Princess Sumaya University for Technology

PSUT

Since 1991 تأسست

جامعة الأميرة سمية للتكنولوجيا

صندوق دعم البحث العلمي

Scientific Research Support Fund

Jordanian Journal of Computers and Information Technology

August 2016 VOLUME 02 NUMBER 02 ISSN 2415 - 1076 (Online)
ISSN 2413 - 9351 (Print)

JJCIT

www.jjcit.org jjcit@psut.edu.jo

An International Peer-Reviewed Scientific Journal
Financed by the Scientific Research Support Fund

Jordanian Journal of Computers and Information Technology (JJCIT)

The Jordanian Journal of Computers and Information Technology (JJCIT) is an international journal that publishes original, high-quality and cutting edge research papers on all aspects and technologies in ICT fields.

JJCIT is hosted by Princess Sumaya University for Technology (PSUT) and supported by the Scientific Research Support Fund in Jordan. Researchers have the right to read, print, distribute, search, download, copy or link to the full text of articles. JJCIT permits reproduction as long as the source is acknowledged.

Aims and Scope

The JJCIT aims to publish the most current developments in the form of original articles and review articles in all areas of Telecommunications, Computer Engineering and Information Technology and make them available to researchers worldwide.

The JJCIT focuses on topics including, but not limited to: Computer Engineering & Communication Networks, Computer Science & Information Systems and Information Technology and Applications.

Indexing

JJCIT is indexed in:

- ScopeMed: www.scopemed.org
- Index Scholar: www.indexscholar.com

Editorial Board Support Team

JJCIT Address

Website: www.jjcit.org

Email: jjcit@psut.edu.jo

Address: Princess Sumaya University for Technology, Khalil Saket Street, Al-Jubaiha.

B.O. Box: 1438 Amman 11941 Jordan.

Telephone: +962-6-5359949.

Fax: +962-6-7295534.

Jordanian Journal of Computers and Information Technology (JJCIT), Vol. 2, No. 2, August 2016.

A SURVEY ON RECENT ADVANCES IN IPTV

Hussein Al-Zoubi[1], Mohammed Halloush[2], Zakaria Al-Qudah[3] and Osameh Al-Kofahi[4]

Department of Computer Engineering, Yarmouk University, Irbid, Jordan
<halzoubi[1],mdhall[2],zakaria.al-qudah[3],osameh[4]>@yu.edu.jo

(Received: 21-Jun.-2015, Revised: 24-Aug.-2015, Accepted: 19-Jan.-2016)

ABSTRACT

After the success of carrying voice over IP networks, the dream of an IP-based television is being turned into reality with the research and standardization efforts. Today, standardization organizations, researchers, as well as commercial companies are putting huge efforts to make Internet Protocol TeleVision (IPTV) successful. The achieved successes of IPTV are due to many factors, including mainly the advanced video coding techniques, the continuously increasing Internet bandwidth that end users have been enjoying and the mature wired and wireless networking protocols and architectures. Although IPTV still faces many challenges, many content providers as well as Internet Service Providers (ISPs) already started to deliver IPTV services to customers. Nevertheless, the way ahead is still long and more efforts are required. This paper surveys a breadth of research areas related to IPTV. It first discusses the general service architecture of IPTV services and diagnostics in these architectures. It then turns the attention to surveying encoding techniques that can serve as enablers for future IPTV especially Scalable Video Coding. Then, various IPTV distribution approaches including peer-to-peer and Content Distribution Networks are discussed. After that, major work on security and privacy concerns in IPTV is tackled. Finally, a discussion of the major work in wireless IPTV services is discussed with focus on LTE-based service offering. We believe such a survey will be very helpful for researchers who would like to educate themselves in the overall landscape of IPTV before digging deeper to address open research problems to bring IPTV closer to reality.

KEYWORDS

IP-based TV, P2P, LTE, SVC, Channel.

1. INTRODUCTION

Although IPTV has been existing for a while, it has recently attracted content and service providers. Part of this is probably due to the exponential growth in technology and the advancements and maturity of networking tools and protocols in addition to multimedia and video compression techniques that have changed many things during the past two decades. These changes rendered IPTV as a viable application. IPTV, therefore, continues to grow commercially.

Three main Internet services are requested by customers today: Internet access, voice over IP (VoIP) and IPTV. The combination of these services is referred to as "Triple-Play" service. Adding mobility to the Triple-Play brings up the "Quadruple-play". Moreover, an IPTV service may deliver both live TV and VOD, where each of these classes has its own characteristics, and is handled differently by IPTV users or clients, which makes the IPTV system design a challenging process.

Two factors have helped IPTV advancement to a great extent: the Advanced Video Coding (AVC) standards; namely MPEG-4 and H.264 and the increased capacities supported by broadband Internet access [2]. AVC standards play an important role in alleviating the effects of the heavy burden that IPTV imposes on the backbone as well as access networks. On the other hand, the ever increasing data rates end users continue to enjoy complements to this role, as these data rates allow more demanding applications to be implemented in practice.

Concerning IPTV, the access network represents the bottleneck. Because recent access technologies, such as Long-Term Evolution (LTE), provide subscribers with high data rates, such technologies are strong candidates to support wireless and mobile IPTV. We, therefore, give LTE special attention in this paper, where we shed light on a number of subjects regarding carrying IPTV traffic over LTE access networks. Many researchers have proposed P2P IPTV architectures to complement IP multi-casting for streaming multi-media and providing TV channels to users. We discuss these proposals as well.

Video transmission over IP networks can be classified into four types [3]: IPTV (the topic of this survey), IPVOD, Internet TV and Internet Video. IPTV delivers content that is similar to broadcast or cable TV, consisting of a huge number of live channels (pre-recorded videos or shows can be delivered too), on a privately owned and operated network (e.g., U-verse by AT&T and FiOS by Verizon). Usually, users are required to subscribe to an IPTV service and own a special device called Set Top Box (or STB for short) to be able to view the content. Similar to IPTV, IPVOD may require a subscription (e.g., Netflix and Amazon) or, alternatively, a per video fee may be required or it can be supported entirely by advertisements. Unlike IPTV, IPVOD is delivered via the Internet, which is a public network. In addition, the delivered content in IPVOD consists of pre-recorded videos, like movies or episodes of shows, and not live channels. Internet TV is similar to IPTV in the type of delivered content, where both live channels and pre-recorded videos can be delivered. However, it is usually a free or an Ad-supported service that does not require an STB to be viewed, since it is delivered via the Internet. In contrast to Internet Video, all the previously mentioned services (IPTV, IPVOD and Internet TV) deliver professionally produced content, whereas the content in Internet Video is usually user generated pre-recorded videos. Internet video is similar to both IPVOD and Internet TV in that it is delivered through the Internet and can be supported by Ads.

With the above classification, we notice that the distinguishing feature of IPTV is that it is delivered on a private network, which has the following advantages:

1. Since IPTV is delivered on a private network, Quality of Service (QoS) can be easily implemented and managed. This is because all network devices are managed by the same operator, which allows for enforcing consistent QoS rules among the whole network. On the contrary, guaranteeing QoS is hard on the public Internet, and thus for IPVOD, Internet TV and Internet Video.
2. True multi-cast can be used, which relies on the Internet Group Management Protocol (IGMP) that allows duplicating packets in-network when they are near their destinations. This reduces the used network resources when compared to multiple unicast sessions, which is used for other types of video services.
3. Real-time Transport Protocol (RTP) over UDP (RTP/UDP) can be used, which is more suitable for multi-casting and streaming video traffic. RTP/UDP is usually not used in the Internet, because of high jitter and the problems encountered in passing firewalls.

It is worth mentioning that despite the above classification, we note that the boundaries between these four types of services are not very firm. That is, some of the available services do not clearly belong to any of the discussed classes; rather, they may borrow from the attributes of more than one class of service.

We identify the following four major areas where IPTV needs further development. First, enhanced security mechanisms with a special focus on user privacy need to be implemented. Second, channel switching delay needs to be efficiently minimized. Third, efficient distribution mechanisms (e.g., peer-to-peer, CDN-based, etc…) should be proposed and fourth, efficient compression mechanisms (especially those based on Scalable Video Coding) need to be devised. These problems become more challenging when considered in the context of real-time IPTV.

The rest of this paper is organized as follows: Section 2 outlines the major recent research directions in IPTV. Then, in Section 3 we talk about wireless IPTV. Section 4 gives the reader an insight into the open research areas and future research directions and concludes the paper.

2. MAJOR RESEARCH DIRECTIONS IN IPTV

In this section, we go over the major recent research directions in IPTV. Extensive research has been performed in the past few years on IPTV. This section provides a brief description of the IPTV service architecture and diagnostics, Scalable Video Coding (SVC), IPTV distribution and security.

2.1 Service Architectures

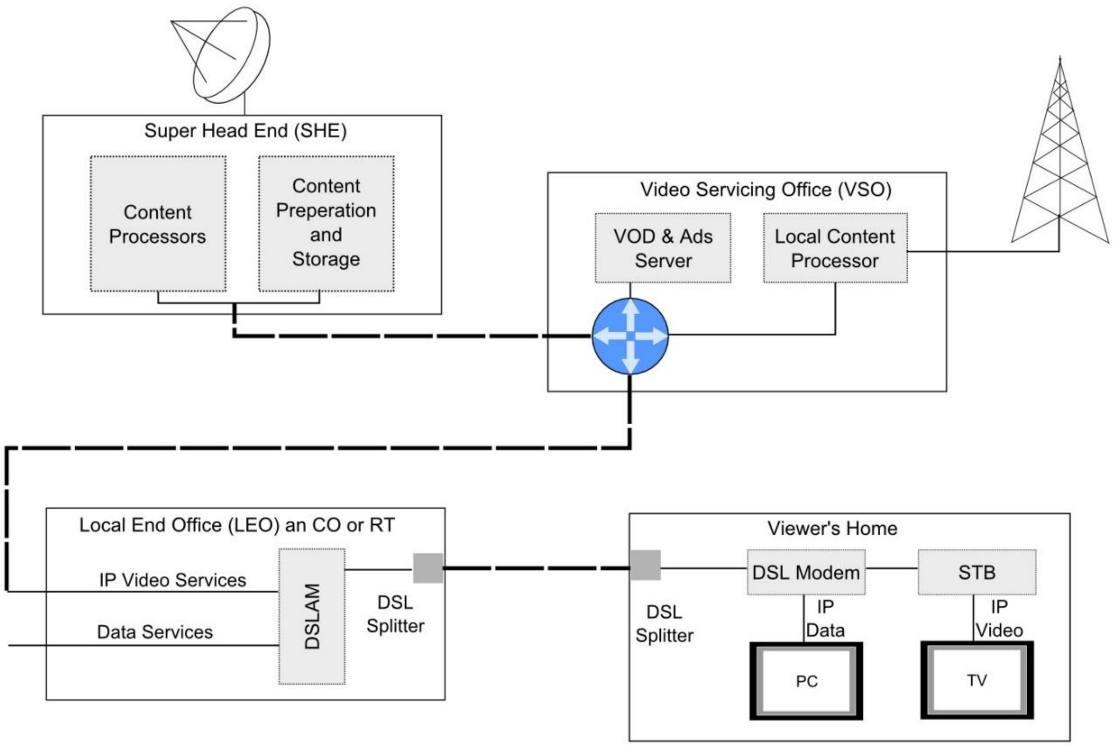


Figure 1. General IPTV Architecture [3].

The general IPTV system architecture is shown in Figure 1 [3]. The Super Head End (SHE) aggregates live content from different programming suppliers, and gathers and formats content intended for VOD (e.g., from videotapes or DVDs). The gathered content is then transcoded into a format suitable for transmission. The SHE labels and catalogues the gathered content and prepares its metadata (description, duration, etc…). After preparing the content, the SHE is responsible for transporting it to the Video Servicing Offices (VSOs). Note that in addition to content, the SHE is capable of processing Ads if they are to be sent to the whole IPTV system. Alternatively, Ads can be handled at the VSOs if they are focused towards a certain region. In addition for being responsible for the processing and the real-time delivery of content in a certain geographical region, the VSO can gather, prepare and deliver local content (e.g., local broadcast channels). The VSO creates IPTV streams and sends them to Central Offices (COs) or Remote Terminals (RTs). Moreover, the VSO handles VOD storage and provides video control capabilities (play, pause, rewind, etc…) for VOD, and interactivity support (making purchases or vote) for live channels. Another important task that the VSO takes care of is STB authorization to make sure that only paying users are receiving the content. Local End Office (LEO) that can be an OC or an RT is simply part of the existing infrastructure installed by telephone companies. It contains equipment to deliver IPTV over DSL lines. Basically, it contains one or more Digital Subscriber Line Access Multiplexer (DSLAM), which is sometimes called Video Remote Access Device (VRAD) that forwards each IPTV stream to its corresponding recipient. DSLAM can also support multi-casting using IGMP, which allows the replication of IPTV streams to end users.

Generally, there are two major service architectures for providing these services. The one-to-many service architecture is used by most of IPTV service providers as discussed previously. In this platform, the service provider specifies for each region one regional station for delivering on-demand IPTV services to subscribers. In other words, each subscriber is served by one pre-specified regional station. This paradigm has many shortcomings in that IPTV service providers are required to replicate devices massively in order to achieve reliability. Moreover, the nature of on-demand services suffers from temporal and spatial skewness, which leads to poor service quality at peak time as well as overall under-utilization of resources. Furthermore, such one-to-many service architecture may miss chances of reducing disk IO or fully utilizing memory-based cache. A collaborative IPTV for delivering on-demand services using many-to-many subscriber-to-station mapping is a viable solution to address these shortcomings. In this many-to-many architecture, a station serves subscribers from multiple regions and one subscriber is served by multiple stations [4].

2.2 Diagnostics in IPTV Architectures

The end-to-end diagnostic mechanisms in IPTV architectures are very useful for improving network performance and reducing customer complaints, which leads to lower operational expenditure (OPEX) costs [7]. The purpose of these diagnostics is to detect configuration problems, including port misconfigurations and element misconfigurations. Diagnostic mechanisms are also used to detect sub-optimal configurations that still allow the network to operate but result in poor network performance. Moreover, diagnostic tools are also useful for addressing congestion due to high traffic volume, server being overloaded and incorrect parameter settings.

The implementation of diagnostic tools can use either in-band or out-of-band mechanisms. In-band-based diagnostics use Operations, Administration and Maintenance (OAM) messages. In-band mechanisms typically collect small amounts of information from many nodes along a path in real time and are therefore considered fast mechanisms. Cross-product end-to-end diagnostics usually use in-band mechanisms. Telnet and Simple Network Management Protocol (SNMP) can be used to communicate diagnostic information out of band. Unlike in-band mechanisms, out-of-band network management system (NMS)-based diagnostics typically collect large amounts of information from fewer nodes and are slower. Table 1 lists some typical symptoms in IPTV architectures, their possible causes and the corresponding diagnostic tool used in each case. All these diagnostic tools may be implemented in-band or out-of-band [7].

Diagnostic tools usually depend on active diagnostic protocols, where some sorts of control messages are injected in the network, then received and processed. On the other hand, passive diagnostic techniques do not require any additional messaging, and thus, no extra overhead. To the best of our knowledge, such techniques are still unexplored for IPTV architectures.

2.3 Scalable Video Coding (SVC)

The scalability of SVC includes temporal, spatial and quality modes. Spatial scalability provides video with reduced picture size, temporal scalability provides video with reduced frame rate and quality scalability (also called Signal-to-Noise Ratio (SNR) scalability) provides video with reduced quality (or lower SNR) [8]. The scalability of SVC thus can lead to several benefits that substantially help in achieving ubiquitous IPTV including efficient methods for graceful degradation in addition to adaptation in both bit rate and format [9]. Moreover, the inherent scalability of SVC will bring many benefits to the wireless domain, since it enables true multi-cast and avoids the more costly solution of multiple unicast sessions. The standardized extension of H.264/MPEG4-AVC that includes scalable video coding specifies three profiles for SVC: Scalable Baseline, Scalable High and Scalable High Intra profiles [8]. The Scalable High profile, which supports spatial scalability (a key feature of SVC), was designed for broadcast and is expected to be used in IPTV [9].

Table 1. Typical symptoms in IPTV architectures, their possible causes and the corresponding diagnostic tool used for each case. All these diagnostic tools may be implemented using in-band or out-of-band mechanisms [7].

Symptoms	**Possible Causes**	**Diagnostic Tool**
A multi-cast channel is not received at home	1. An Internet Group Management Protocol (IGMP) join issued by the user for a given channel was dropped due to overflow. 2. The IGMP proxy function did not work. 3. The multi-cast forwarding database was updated incorrectly. 4. The multi-cast forwarding database was overflowed.	Multi-cast link trace (MLINK)
Lack of connectivity to a given channel	The user may not be allowed to access the network resources. This includes bandwidth availability and current bandwidth usage for a given channel and a specific stream type.	Information access line (iLINE)
Degradation of quality for a certain service	Alteration of packet priorities. Priority misconfigurations will result in an operational but sub-optimal network that may not trigger an action from network administrators. Priority misconfiguration may cause only transient problems and can remain undetected for a long time.	Priority trace (pTrace)
A user loses connectivity to some or all IP services	Dynamic Host Configuration Protocol (DHCP) misconfiguration.	In-band (lease query) mechanisms or NMS-based Diagnostics

SVC supports the transmission of video in layers; a base and one or more enhancement layers. Users receiving the base layer can decode and display video of base quality. Decoding more layers enhances video quality. This allows video transmission to receivers with different capabilities where each receiver can decode a number of video layers (based on its needs and capabilities) from a video stream that has several layers. This is more efficient than sending different video streams to meet the different receivers' needs.

The aforementioned features of SVC can be very helpful for IPTV. Specifically, applying SVC in the context of IPTV can significantly improve content portability, optimize content management and distribution, achieve a smart management of access network throughput and improve QoS and QoE [9]. The scalability of SVC helps in improving content portability in that the user can watch a video content on devices with different capabilities like TV, PC, smart phone, etc… even if this video content was not downloaded on that device, keeping in mind that these various devices differ from each other in terms of display area, processing capabilities, mobility and power consumption. In terms of optimizing content management and distribution, the scalability of SVC serves different subscribers at the best of their connection bandwidth (for example, half HD or full HD). Moreover, the scalability of SVC helps in managing the throughput of the access network in a smart way, because the adaptation of SVC simply drops selected IP packets and adapts the video bit rate to achieve the suitable user experience. SVC plays an important role in improving QoS and QoE by reducing channel zapping time and improving content navigation in the case of VoD IPTV. In addition, the graceful degradation feature of SVC aids in congestion control, as well as in handling access network constraints [9].

The architecture of SVC in H.264/SVC consists of several layers, where the original H.264/AVC video stream is called the base layer. The scalability of SVC is achieved by adding additional enhancement layers to the base layer for the purpose of increasing frame rate, picture size and quality (SNR). H.264/SVC employs inserting Instantaneous Decoding Refresh (IDR) frames into the video stream in order to enable random access. The simulcast technique [10] works by including a companion stream that has more IDR frames at lower quality as illustrated in Figure 2 [10], where black rectangles represent IDR frames and arrows represent inter-prediction. The main video stream shown in Figure 2 (a) has full quality and slow switching, while in Figure 2 (b) it has lower quality and faster switching. The two video streams in IPTV are independently encoded and independently transmitted. The low-quality companion stream works as channel-switching acceleration in that the screen of the end user when he/she changes the channel is first displayed. The regular SVC configuration [10] combines in one scalable video stream both the main and the companion streams, where the companion stream is encoded as the base layer and the main video stream is encoded as the enhancement layer. This regular SVC configuration can reduce bandwidth in the core network. However, it puts a burden on the access network in terms of bandwidth requirements, because the base layer is still to be transmitted to the user's STB if no channel change occurs.

Because of impairments and network conditions like congestion, packets of IPTV are liable to errors and loss, which badly influences the video quality. Error concealment for SVC in IPTV systems is therefore necessary. The four error concealment methods that have been adopted for SVC are: temporal direct, frame copy, motion and residual upsampling, as well as base layer reconstruction and upsampling [11]. The first two methods are intra-layer error concealment methods, while the last two are inter-layer error concealment methods. Implementing error concealment techniques at the decoder increases the decoder's complexity [12]. The work in [3] proposes a per-packet prioritization mechanism that assigns a high priority to important frames, such as IDR frames, P-frames and B-frames. This mechanism provides resiliency against transmission errors for such important frames.

Other related research topics include proposing resource management mechanisms for scalable transmission control to optimize resource allocation for SVC-based mobile IPTV service [13]. The proposed system properly trades resource consumption and utility gain and maintains a given level of service coverage for each layer of the scalable video-encoded stream.

2.4 IPTV Distribution

To provide customers with IPTV Services with High QoS, there are two different architectures for the distribution of servers in the network [15]: centralized and distributed. The pros and cons for each approach are summarized in Table 2.

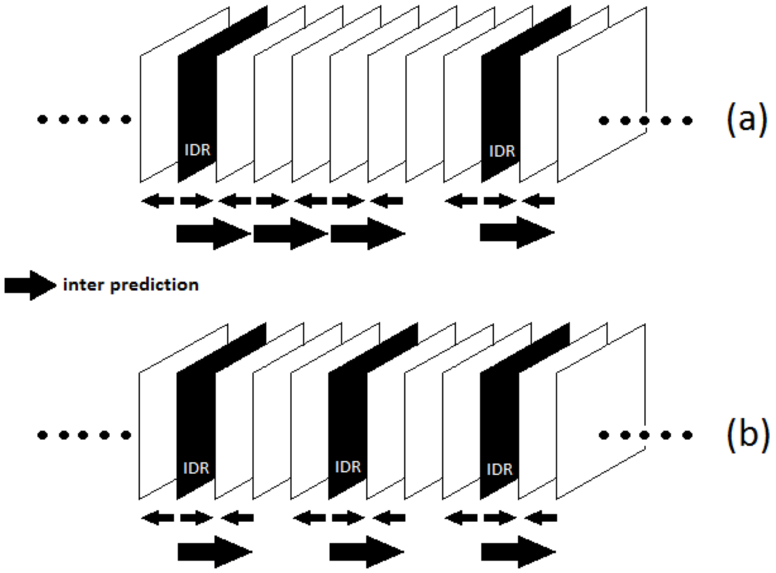


Figure 2. The simulcast of H.264/SVC: (a) Main video stream, (b) Companion video stream.

2.4.1 P2P IPTV

It is expected that P2P IPTV will be more prevalent than client-server IPTV in the future. The main reason for this is that P2P scales better. Currently, P2P IPTV is only deployed in China. UUSee Inc. [66] is a principal P2P live streaming provider in China [67]. Nonetheless, intensive research has been conducted on this topic. In P2P IPTV, peers contribute their resources such as CPU and upload bandwidth and at the same time upload (they act as content providers) and download real-time video streams, which adds no additional infrastructure cost as the community grows larger. This has been reflected in practice, where it has been noticed that the traffic generated by P2P IPTV applications has recently increased significantly [16]. A similar conclusion has been drawn in [17], where the authors argue that IP multi-casting would be the best solution for streaming TV content of popular TV channels that are greater in number than the active multi-cast groups. However, for less popular channels and using unicast connections, P2P produces comparable performance in terms of scalability, bandwidth utilization and quality. The authors in their study took into account the findings in [18] which show that the popularity of TV channels follows the Pareto principle, where only a small portion of people who watch TV channels are interested in watching the vast majority of TV channels. Similarly, picking the end-to-end content blocking as a QoE parameter, P2PTV outperforms the IPTV as the total number of users increases [19].

P2P IPTV was originally proposed as a solution for the scalability issue of streaming multi-media content at lower costs. The scalability for video distribution using the P2P approach is achieved by dividing the video into several chunks prior to distribution. These chunks are then shared among peers [20]. In P2P IPTV, an overlay network is created for each channel to connect the viewers (peers) of that channel, who receive, view and redistribute the content of the channel [21]. Some of the most famous IPTV services that employ P2P video streaming include PPLive, PPStream, SOPCast and TVants. However, there are several concerns regarding the success of P2P IPTV, including the stress that P2P imposes on ISP networks, because of the overwhelming traffic that P2P may pump into core and access networks, in addition to security concerns as it opens the appetite for malicious attacks [22]. Moreover, one of the issues that has been considered as a limitation of the P2P applications is having users (peers) being inside Network Address Translated (NATed) networks. This is an issue that can be faced by any commercial IPTV service provider. The NAT issue has been investigated in [23] with two categories of peers: the first category includes peers inside Lancaster University, the majority of whom are behind NAT. The second category represents off-campus peers, who had poor sharing and in general performed poorly, because they were unable to contact their on-campus peers and therefore started to starve of content quickly. The authors were able to resolve this problem by increasing the piece availability to external peers.

There are two different approaches for streaming IPTV using P2P architectures: the mesh-pull approach and the tree-push approach, where the mesh-pull succeeded in achieving commercial deployments. The P2P IPTV architecture followed in [24] for IP Multi-media Subsystem (IMS) employs a centralized control layer in addition to a P2P-like media layer at the peers, who act as forwarding nodes of multi-media streaming. The architecture in [25] integrates P2P distributed systems into Telco's STB at home to support advanced rewind functionalities and large number of live channels. Bikfalvi et al. [17] proposed a hybrid IPTV system that combines IP multi-casting and a P2P overlay unicasting to provide TV channels to large number of receivers as shown in Figure 3 [17].

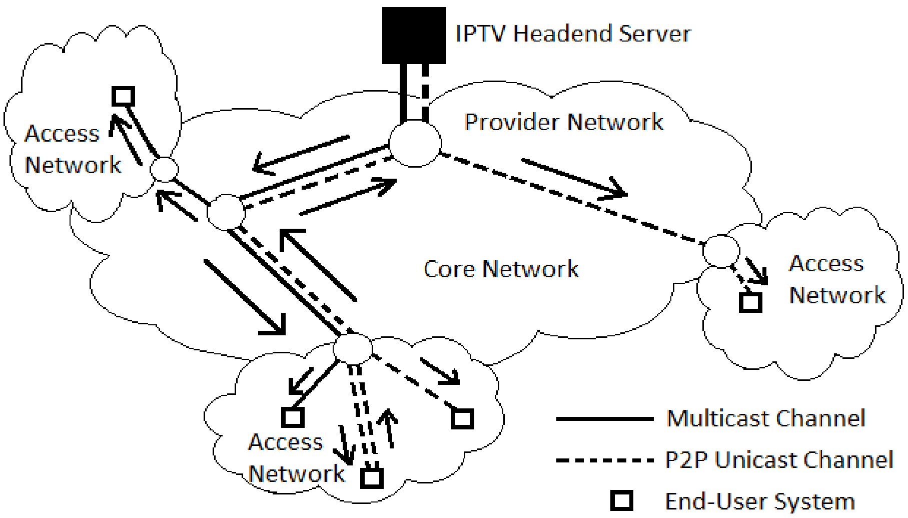


Figure 3. Multi-cast and unicast hybrid IPTV systems [17].

Decreasing channel zapping is one of the most important challenges in IPTV in general and P2P IPTV in particular. When employing P2P networks for IPTV, a delay is incurred when switching from one channel to another. Each IPTV channel is an overlay network. Switching from one channel to another is done by moving to another overlay network. This transition incurs a delay that ranges from seconds to minutes.

As a solution to the channel zapping problem, other channels' (adjacent channels') data can be sent along with the current channel data. When the user switches to an adjacent channel, the switching occurs with no delay [26]. In [27], another solution to the channel zapping problem is provided by sending I frames of the adjacent channels along with the video frames of the current channel, so that when the user switches to an adjacent channel the additional I frames are used instead of waiting for the normal I frames [27].

The channels' data transmitted with the current channel are supposed to have high probability for being selected by the user. Adjacent channels are selected in such a way that the likelihood that the user switches to theses channel is high. Depending on the popularity of channels (less delay incurred when switching to a popular channel than that incurred when switching to a less popular channel) [28].

Hei *et* al. [29] conducted a measurement study of PPLive and concluded that the internet infrastructure is sufficient for a large scale IPTV system. On the other hand, there are some P2P specific issues that need to be resolved, so that a P2P IPTV system becomes more attractive. Slow start due to startup delay caused by tens of seconds of buffering before playback helps overcome the fluctuation of the video connections' rates. Also, dedicated infrastructure may become necessary to support higher data rates at peers that have upload capabilities that are less than the download capabilities. The dedicated infrastructure consists of nodes that have high upload capacity. Variance in playback times at the different peers is another issue that needs to be considered.

Other P2P IPTV related topics include providing peer selection algorithms that offer reliable media streaming transmission as well as satisfactory user QoE [30], studying the overlay characteristics of large-scale P2P IPTV systems and identifying differences between PPLive multi-media streaming overlays and P2P file-sharing overlays [31], providing channel zapping (surfing) distributed mechanisms to reduce channel zapping time in P2P IPTV that involve switching to another P2P overlay network [21], as well as designing data scheduling schemes for storage-constrained mesh-based P2P IPTV systems [32].

Despite all the research conducted on P2P IPTV, there are still many challenges that need to be addressed, especially as prices of infrastructure CDN-based IPTV is getting lower [33]. The inherent heterogeneity of the Internet and user nodes makes it impossible to find a solution that would be suitable for all cases. Some of the main challenges that face P2P IPTV are fairness, availability, stability and including incentives in a P2P IPTV system (tit-for-tat is not suitable, especially for real-time IPTV). Also, one cumbersome situation when it comes to P2P IPTV is dealing with flash crowds, where a huge number of users become interested in an event all at once (e.g., a live event) [33].

2.4.2 VN-based Networks for Delivering IPTV Services

The current solutions for delivering IPTV services is by using a backbone or IP overlay-based content network. However, these solutions suffer from limitations in terms of guaranteed service delivery, cost effectiveness, flexible control and a scalable network infrastructure. The virtual network (VN) is a promising alternative, which allows multiple service providers to share the infrastructure by dividing the physical network structure of infrastructure providers (InPs) into several slices and associating them with different service providers. The main advantages of applying VN for delivering IP services can be summarized in the flexibility and control that VN offers to IPTV service providers in addition to security and reliability, since multiple VNs are isolated from each other even if they coexist on a shared substrate network [34]. Delivering IPTV over VNs requires leasing servers and links from an infrastructure provider and then replicating content according to the demand. The work in [35] integrates these two problems and proposes a genetic algorithm as a heuristic.

There are, however, two major challenges that must be worked out when considering VN for IPTV service delivery [34]. The first is the need for a well-designed virtual network topology (VNT) to ensure that services can be efficiently delivered from the servers to the customers. VNT must be efficient, cost effective and reliable to overcome any virtual node or link failure. The second challenge is the handling of a VN allocation failure problem by infrastructure providers using an efficient VNT reconstruction algorithm.

2.4.3 Content Distribution Networks (CDNs)

Commercial video distribution services (VoD and Live IPTV) over CDNs have been around for a while, like Akamai video streaming service and Limelight streaming service. In addition, Netflix, Amazon and AppleTV are examples of commercial IPTV service providers assisted by CDNs. The massive, highly distributed infrastructure that CDNs own enables them to serve the video content from a lightly loaded location nearby end-users. Since these commercial CDNs deploy sophisticated measurement and quality path selection mechanisms, they are able to avoid points of congestion in the network, and therefore, provide high quality streaming services. Furthermore, video distribution over CDNs eases the burden on the core of the network posed by the high bandwidth demand of the IPTV application. In terms of research, Yin et al. [36] proposed a hybrid architecture that combines P2P distribution with CDN-based distribution. The proposed architecture benefits from the reliability and quality of CDNs and from the inherent scalability features of P2P systems when the total demand exceeds the capacity of the CDN. The authors also reported on the performance of a commercial hybrid CDN and P2P IPTV service. The authors in [37] showed that the unique characteristics of video transport; namely the high bandwidth requirement and the design of Akamai's CDN streaming service can leave the system susceptible to intentional service degradation (by an attacker). Therefore, the authors demonstrated the importance of careful design of these services to avoid possible attacks.

2.5 Security and Privacy

When considering security in IPTV, many important issues should be considered. The authors of [42] argued that to let end users enjoy content in new ways without any security hurdle, IPTV security must be architected and implemented correctly. This is possible by providing a cost-effective security and allowing service providers to explore and innovate new business models, as well as customer experiences. There are a number of security requirements for the transmission of any form of data on the Internet in general and for the transmission of IPTV traffic in specific. These requirements include the following [43]:

1. Confidentiality, which means to allow legitimate users only to recognize the transmitted data and to prevent attackers from getting such information as the source of data, destination, time, length and traffic characteristics of communication channels. To achieve confidentiality, the transmitted data must be encrypted.
2. Integrity; that is to protect data, either stored on an information system or being transmitted, from being falsified. To achieve integrity and notice any illegal data modification, electronic signatures can be employed.
3. Authentication, where during the transmission of data, the sender should be able to confirm the identity of the receiver and the receiver should be able to confirm the identity of the sender. Authentication is in general hard to achieve. A form of user ID and password can however be used.
4. Access control, which means to grant access to data and network resources to authorized users only. To control access to network services, intrusion detection systems and firewalls can be used to counter attacks of unauthorized users.

Authentication, Authorization and Accounting (AAA) is a standard developed by the IETF working group to provide secure network access, user authentication and accounting [43]. The Conditional Access System (CAS) is used to control users' access to charged broadcasting services. DVB CAS [44] is the standard developed in Europe and ATSC CAS in the USA [43]. The Digital Right Management (DRM) is used in Internet-based environments to manage intelligent property of digital contents [43]. The time-constrained management rights provided by DRM include the right to read, distribute, edit and copy [45].

Content protection is required to enable subscribers to use the content they have acquired in accordance with the rights they have been granted. To protect content and services of IPTV, a number of security aspects need to be addressed, including identification, authentication, authorization, key distribution, content encryption, right's expression, renewability and revocation, content export, compliance and robustness [46].

Possible security attacks in IPTV include subscriber and STB impersonation attacks, replay attack, stolen verifier attack, smart card loss attack, man-in-the-middle attack and attack on perfect forward secrecy [47]. Moreover, providing IPTV through IP networks may result in illegal control, illegal content distribution, service theft, access of unapproved users, sniffing, tapping, Denial of Service (DoS) attack, War Dialling attack, Rogue Device attack and harmful software infection [43].

Secure IP Multi-cast allows the secure transmission of IPTV services to groups of receivers. Nevertheless, when users switch between groups (channel zapping situations), the existing solutions do not optimize the signalling generated by the IPTV, because new cryptographic material is retrieved by the zapping user [48]. Moreover, Secure IP Multi-cast neglects access control and network management [49]. IPTV services use one key for each video channel. For perfect forward and backward secrecy, key distribution for secure group communications applies key refreshing techniques when group members join and leave (when a group changes) [49].

Subscriber authentication schemes for mobile IPTV users require individual-level user authentication to provide personalized, tailored services for mobile users. These authentication schemes are different from the existing IPTV authentication schemes that provide the same IPTV services and access levels to the whole family members using a set-top box (STB)-level authentication [50]. Secure key exchange with mutual authentication between STB and smart card is required to avoid smart card cloning and McCormac Hack attacks [47].

The provision of secure authentication mechanisms is a fundamental security requirement of IPTV. The traditional protection scheme's like Conditional Access System (CAS) [51] and Digital Rights Management (DRM), may fail in IPTV, because IPTV provides interactive on-demand services [50]. Moreover, mobile IPTV requires individual-level user authentication in contrast to the authentication schemes that are currently used in IPTV and provide the same services to the whole family members using a set-top box (STB). Existing IPTV authentication schemes are either password-based, RFID-based or Universal Subscriber Identity Module (USIM)-based [50].

According to the ITU-T IPTV Focus Group [52], there are five types of security threats facing IPTV:

1. Content security threats, which include interception, unauthorized viewing and unauthorized reproduction or redistribution.
2. Service security threats, which include violating copyrights of the programs which IPTV service platform provides to the subscribers, masquerading/spoofing IPTV service provider, malicious threats aimed at the IPTV servers and theft of the subscribers' information.
3. Network security threats, which include intentional threats to the network equipment or resources, security threats to multi-cast techniques used in IPTV bearer network and malicious attacks on nodes in content distribution network.
4. Terminal device security threats, which include illegally accessing clear content by tampering device hardware or software, illegally accessing keys or other secret information in devices using software cracking or hardware tampering, device malfunctioning by software and hardware methods, unauthorized applications, the failure of terminal equipment, unauthenticated terminal devices connecting to the home network and unauthorized use by subscribers.
5. Subscriber security threats, which include theft of the subscribers' information and end user's privacy by malicious programs.

These types of security threats are illustrated by the Security Threats Model shown in Figure 4 below, which also illustrates the relationships between these threats [52].

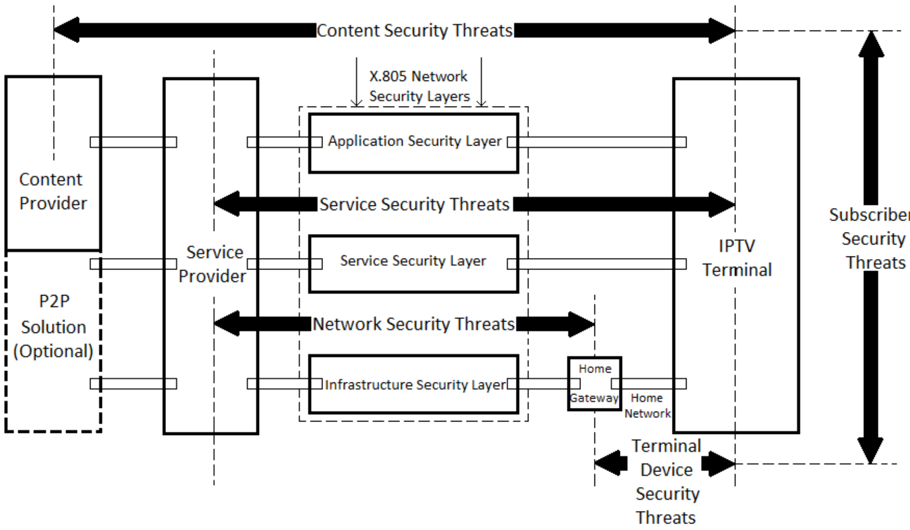


Figure 4. Security threats model [52].

Privacy protection is another important security aspect that must be enforced in order to achieve a secure IPTV. The recommender service is employed by service providers to attract and satisfy customers by collecting information about the user preferences and selecting the items that might be of interest to customers in the future. Currently, the recommender services are implemented in a centralized fashion using a collaborative filtering technique, a smart machine learning technique that can recommend new items for users by collecting preference data for all the users in one place [53]. This, of course, may violate the privacy of users.

3. IPTV Over Wireless Networks

IPTV services can make use of the wireless broadcast advantage. Unlike transmissions on wired networks, wireless transmissions are broadcast by nature. Therefore, an IPTV client actually receives the channel he/she requested in addition to all other channels broadcast in his/her proximity to neighbouring clients. This characteristic can help in improving the channel zap times as suggested in [54] for LTE-based networks.

Recently, multi-Gigabit wireless personal area networks (WPANs) utilize the 60 GHz millimeter-wave (mmWave) communication technologies. For mmWave-based WPANs, directional antennas are preferred over omni-directional antennas. This is because of the small wavelength of mmWave communication systems. The attenuation of mmWave signals in the air is much faster than lower frequency signals. Furthermore, it is possible to concurrently allow P2P transmission and exploit spatial multiplexing using directional antennas [55]. This has the potential of guaranteeing the QoS requirements of IPTV and improving resource utilization efficiency.

3.1 Mobile IPTV

The H.264/MPEG4-AVC video codec has been implemented in STBs, Flash and QuickTime players, and is expected to be included in mobile devices [9]. Scalable High Profile of SVC will not probably be supported in mobile IPTV, because interlaced coding is not used in the mobile space due to the high implementation complexity [9].

ITU-T classifies IPTV architectures into Next-Generation Network (NGN)-based and non-NGN-based architectures [36]. Figure 5 shows a typical mobile IPTV architecture, where both IPTV sender and receiver can be mobile. Moreover, mobile IPTV users are allowed to provide content to other mobile and non-mobile IPTV users. The network architecture could be non-NGN, NGN IMS or NGN non-IMS based.

Point-to-Point (PTP) is usually employed in mobile communications. Unfortunately, PTP may quickly cause exhaustion in the network resources represented mainly by the bandwidth. Therefore, to restrict the limitations imposed by the bottleneck that exists at the last mile from the mobile grid backbone in terms of data rates and the amount of communication, Point-To-Multipoint (PTM) using broadcasting and multi-casting techniques is advisable, putting in mind that interests in certain TV contents like sports and news is common by many mobile users [36].

In a typical situation, it is possible that many home devices request IPTV services simultaneously. Using a single Access Point (AP) will not be adequate to provide a full coverage at home. One possible solution is to use several APs, which might lead to interference among devices at the same house and within neighbourhood. This will result in more collisions and increase the packet loss ratio to an unacceptable value. To manage the wireless transmission of IPTV services at home, two possible approaches can be used: distributed and centralized architectures. Since home networks are centralized by nature (home network is connected to ISP via a single gateway), the centralized approach seems more appropriate for home networks [36]. In light of this, several commercial products implementing the centralized architecture are available today, including OmniAccess series from Alcatel-Lucent, RFS series from Motorola, Lightweight AP and WLC from Cisco, just to mention some.

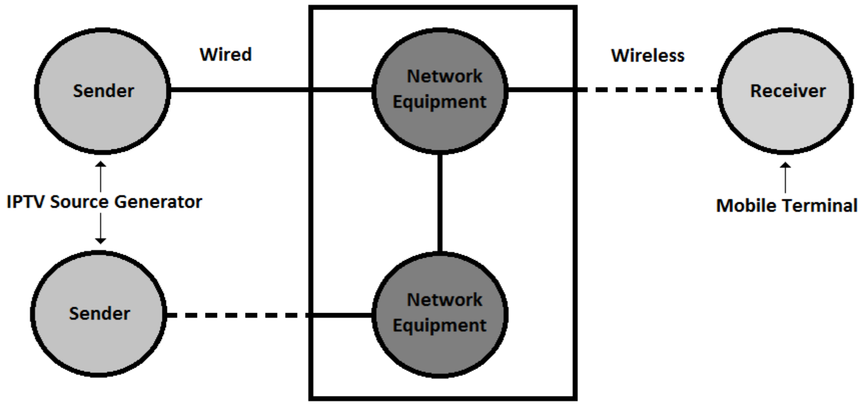


Figure 5. A general architecture of mobile IPTV.

3.2 Long Term Evolution (LTE)

LTE is the fourth-generation cellular network that has proven to provide high-speed wireless and mobile Internet. Multi-media multi-casting, streaming and downloading in LTE networks are provided through evolved Multi-media Broadcast and Multi-cast Services (eMBMS) [56]. Using LTE for live broadcast (live events or TV channels) is not available commercially yet. The closest LTE got to being used for live broadcast is in a few trials to broadcast in-stadium live events to the audience in the same stadium [57]. Some service providers, e.g., Verizon (USA) and Telstra (Australia) [58], are planning to use LTE for video broadcast in the near future, but this is not done yet, since the market is favouring OTT content [59]. Ericsson is the first to propose a complete solution to broadcast video over LTE, and in its vision, this will be enabled by HEVC (High Efficiency Video Coding), MPEG DASH (Dynamic Adaptive Streaming over HTTP (DASH)) and eMBMS [60]. Nevertheless, LTE should be capable of serving VOD due to its high supported rates. For example, the minimum bandwidth requirement for Netflix is 0.5 Mbps, while the recommended bandwidth by Netflix, for HD and Ultra HD, is 5Mbps and 25Mbps, respectively [61], which should be satisfied by LTE.

It is important to assess the performance of LTE and to measure the amount of success that has been achieved in the deployment and use of LTE networks. In a recent study, the authors of [62] performed a thorough performance evaluation of LTE and reached a number of interesting findings. The LTE network topology is shown in Figure 6 [62]. The typical LTE network shown in Figure 6 shows that the end users use User Equipments (UEs) to get access to the 4G LTE. The Radio Access Network (RAN) connects UEs to the Core Network (CN). The RAN consists of a number of Evolved Node B (eNB) or base stations. The authors collected data at the Monitor. The role of Performance Enhancing Proxy (PEP) is to split the end-to-end TCP connections between the UE and the server port 80 or 8080 into two connections: the first is between the UE and the PEP and the second is between the PEP and the server. The split is transparent to the UEs. The PEP compresses data and acts as a cache, which leads to improved performance.

LTE has been attracting more users because of the significantly higher bandwidths and lower delays compared to the 3G networks. Nevertheless, the authors of [62] found that many TCP connections underutilize the available bandwidth significantly, which incurs additional energy consumption and makes data downloads take longer delays. The authors found that the main cause of this underutilization of bandwidth is that some applications are energy and LTE network unfriendly. The authors also found that the majority of LTE traffic (in terms of byte count) is multi-media (video, audio and graphics). They found that user consumption of multi-media and the dominance of video have been significantly increasing compared to previous studies. They also found that LTE outperforms WiFi in many cases and substantially outperforms WiMAX.

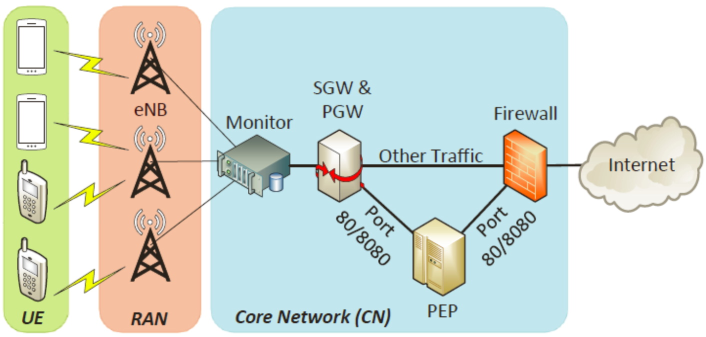


Figure 6. The LTE network topology used measurements in [62].

Employing SVC can be very helpful for IPTV over LTE networks. SVC proved to be successful in reducing packet loss and bandwidth requirements of video multi-casting in LTE with acceptable video quality. SVC also provides graceful degradation of video quality for users in the cell edge, where users near the base station are served with all the three temporal layers (containing I, P and B frames) and the base temporal layer (containing I and P frames) is only sent to users far from the base station [63]. Authors showed 18% reduction in bit rate with SVC compared to simulcasting using H.264 without SVC. In [64], the authors proposed a dynamic AMC and scheduling approach for MBSFN-based SVC video transmission in LTE networks. The authors showed that the proposed scheduling can considerably save radio spectrum.

Power is an important issue in the transmission of video over LTE networks; lower transmission rates help increasing sleep mode power conservation [56]. The capacity of poor receivers affects the eMBMS performance. Existing power settings increase the transmission power for the purpose of better covering poor nodes, which leads to more power consumption and interference. There are two approaches to reduce power consumption in LTE. The first approach multi-casts multi-media using layered video (i.e., through SVC), where the base layer is transmitted using high power at low rate and the enhancement layers are transmitted using lower power and higher transmission rates. The second approach depends on power-efficient multi-casting in the various layers like cooperative and opportunistic routing (network layer), scheduling (MAC layer) and multi-cast beamforming (physical layer). Therefore, power efficient SVC multi-casting should consider power savings in group, channel assign and schedule [56].

For future research, these subjects must be profoundly studied in the case of mobile users in LTE networks, as mobility can cause substantial CQI fluctuations. More efforts are also required to explore the interactions between radio, network transport and application layers. Moreover, new energy-efficient resource management policies and transport protocols are needed. Finally, it is critical to optimize the LTE network for IPTV.

4. Conclusions and Future Research Directions

The increased data rates provided to end users and overwhelming the world with numerous smart devices, like mobile phones and other handheld devices that can access the Internet from anywhere at any time, have turned providing television over IP networks into reality. There are, however, many unsolved issues in the realm of IPTV, especially if it is required to support wireless and mobile IPTV. The complexity and challenges increase while integrating High Definition TeleVision (HDTV) into IPTV. In this paper, we provided an extensive review for the recent research efforts in the field of IPTV on areas like IPTV services, P2P IPTV, SVC and IPTV, security and wireless and mobile IPTV.

Despite all the advancement, in video coding techniques, we believe further study is still needed on how to efficiently encode real-time video because of the SVC's encoder complexity [14]. SVC is a promising solution for providing TV services over IP networks [8]. The technical complexities of the design of SVC features for the consumer end devices will decrease over time allowing commercial deployment.

The work in [4] addressed three issues; request dispatching, content placement and controlling collaboration scope to ensure QoS. We believe that there are still other issues that need to be addressed in many-to-many service architectures. For example, how would such a system perform with real-time live traffic? Also, although network coding [5] has been applied to P2P IPTV (e.g., [6]), the many-to-many service architecture is different. Utilizing network coding in such many-to-many service architectures, which have replicated redundant data, may bring throughput, reliability and cost advantages. Finally, as suggested in [4], algorithms for high definition video delivery in many-to-many architectures still need to be evaluated.

With respect to CDN-based IPTV distribution, there are three main challenges. First, the need to optimize CDN deployment to reduce cost will always be an important challenge. This is due to the continuously increasing demand on IPTV, intense competition in the content delivery market and to the constant need to upgrade and maintain network hardware. The second challenge is end-user miss-location, which is mainly caused by assigning end-users to CDN servers based on the location of the DNS resolver that served the user, which in many cases gives an inaccurate view on user location. Third, since CDNs have limited network knowledge, inappropriate user assignment may lead to overloading network bottlenecks [38]. The above challenges can be addressed through the collaboration between ISPs and CDNs, where [38] and [39] are first steps towards this approach.

In addition, energy-efficient IPTV or Green IPTV should receive more attention. Current IPTV providers' multi-cast TV channels as close as possible to the user are needed to reduce the delay perceived by the user as much as possible. However, not all channels have viewers all the time. Therefore, selective multi-casting channels may reduce the consumed energy, especially for high definition channels [40].

Another research direction that needs further exploration is implementing IPTV services over Information Centric Networks (ICN). ICN-based IPTV is yet to be explored because of the challenging nature of ICN architectures [41].

User's privacy remains a big issue in IPTV. A possible future direction to privacy preserving IPTV as suggested by [53] could include utilizing game theory in the implementation of recommender service.

The most important aspects of the IPTV security system remain the need to provide cost effective state-of-the-art security with an ability to recover from a compromised system and the need to provide a competitive open and interoperable security infrastructure with flexible capability to allow service providers to innovate and explore new business models and consumer experiences. Furthermore, while [53] has proposed a privacy-preserving multi-agent based recommender service, the security of such services needs to be studied carefully. For example, a breach of security of these services may render the service useless (by for example submitting a large number of fake user ratings) or may defeat the goal of these services by being able to trace a rating back to the originating user.

The other open issue that needs to be addressed as a research direction in IPTV security to achieve content service protection [65] is transcodable encryption. That is to decrypt the protected content and re-encode the decrypted content without considerable performance overhead at the end user side.

To reduce the huge bandwidth requirements of multi-media contents, video compression techniques must be employed. Since these compression techniques are lossy in nature, degradation in the perceptual video quality is possible. Therefore, more efficient error control mechanisms are required for better QoS and QoE in IPTV. It is also required to implement error concealment techniques that do not increase the decoder's complexity [12].

Channel zapping remains an issue that must be dealt with for successful IPTV deployment. Channel zapping time should be as short as possible. If the channel zapping time is too long, the quality of experience will be affected, which leads to user dissatisfaction. It is required to come up with fast channel switching SVC configurations that reduce bandwidth on the access network and facilitate backward compatibility [10].

Following the work of [9], we believe that more effort should be put to implement Context-Adaptive Binary Arithmetic Coding (CABAC) and EB slices decoding into mobile devices [9]. Additionally, the devices available today for providing wireless IPTV services are relatively expensive, which makes them not affordable by home users in most cases. Therefore, more research is required to address this issue. Other related topics include dynamic scheduling algorithms [37], as well as rateless channel coding and data-partitioning [36].

Concerning mobile IPTV, future research directions include the derivation of MPLS-based NGN architectures to support QoS, traffic engineering and VPN IPTV mobility requirements, correlating networking resources, terminal capability, and user profile for mobile IPTV streaming applications, integrating SVC coding technologies into mobile IPTV services and devising display procedures for mobile IPTV users that are both viewer centric and context aware [37].

We believe that an evaluation of current service deployments is necessary. It is also important to research new encoding schemes for wireless devices exploiting SVC. Finally, we argue that unlimited data plans might not be affordable now, but it is expected to be affordable in the future, which permits the utilization of the full power of IPTV over LTE networks.

REFERENCES

[1] Ning Liu, Huajie Cui, S.-H. Gary Chan, Zhipeng Chen and Yirong Zhuang, "Dissecting User Behaviours for a Simultaneous Live and VoD IPTV System," ACM Trans. Multimedia Comput. Commun. Appl., vol. 10, no. 3, pp. 23:1-23:16, April 2014.

[2] She, J., Fen Hou, Pin-han Ho and Liang-Liang Xie, "IPTV over WiMAX: Key Success Factors, Challenges and Solutions [Advances in Mobile Multimedia]," IEEE Communications Magazine, vol. 45, no. 8, pp. 87-93, August 2007.

[3] Simpson, Wes and Howard Greenfield, IPTV and Internet Video: Expanding the Reach of Television Broadcasting, CRC Press, 2012.

[4] Shicong Meng, Ling Liu and Jianwei Yin, "A Collaborative and Scalable Platform for On-Demand IPTV Services," IEEE Transactions on Services Computing, vol. 6, no. 3, pp. 358-372, July-Sept. 2013.

[5] R. Ahlswede, Cai Ning, S.-Y.R. Li and R.W. Yeung, "Network Information Flow," IEEE Transactions on Information Theory, Services Computing, vol. 46, no. 4, pp. 1204-1216, July 2000.

[6] Chuan Wu, Baochun Li and Zongpeng Li, "Dynamic Bandwidth Auctions in Multioverlay P2P Streaming with Network Coding," IEEE Transactions on Parallel and Distributed Systems, Services Computing, vol. 19, no. 6, pp. 806-820, June 2008.

[7] Kamakshi Sridhar, Gerard Damm and Hakki C Cankaya, "End-to-end diagnostics in IPTV architectures," Bell Labs Technical Journal, vol. 13, no. 1, pp. 29-34, Spring 2008.

[8] H. Schwarz, D. Marpe and T. Wiegand, "Overview of the Scalable Video Coding Extension of the H.264/AVC Standard," IEEE Transactions on Circuits and Systems for Video Technology, vol. 17, no. 9, pp. 1103-1120, Sept. 2007.

[9] T. Wiegand, L. Noblet and F. Rovati, "Scalable Video Coding for IPTV Services," IEEE Transactions on Broadcasting, vol. 55, no. 2, pp. 527-538, June 2009.

[10] G. Van Wallendael, W. Van Lancker, J. De Cock, P. Lambert, J.-F. Macq and R. Van de Walle, "Fast Channel Switching Based on SVC in IPTV Environments," IEEE Transactions on Broadcasting, vol. 58, no. 1, pp. 57-65, March 2012.

[11] Y. Chen, K. Xie, F. Zhang, P. Pandit and J. Boyce, "Frame Loss Error Concealment for SVC," A Journal of Zhejiang University Science, vol. 7, no. 5, pp. 677-683, 2006.

[12] P. Lambert, P. Debevere, J. Cock, J.-F. Macq, N. Degrande, D. Vleeschauwer and R. Walle, "Real-time Error Concealing Bitstream Adaptation Methods for SVC in IPTV Systems," IEEE Internet Computing, vol. 13, no. 3, pp. 23-31, May-June 2009.

[13] Y. I. Choi, J. W. Kim, J. H. Kim, J. S. Jeong, W. J. Chun and C. G. Kang, "A Resource Management Framework for Mobile IPTV Service with Scalable Video Coding," Proc. of the 4th International Conference on Communications and Electronics (ICCE), pp. 136-140, 2012.

[14] Soohong Park and Seong-Ho Jeong, "Mobile IPTV: Approaches, Challenges, Standards, and QoS Support," IEEE Internet Computing, vol. 13, no. 3, pp. 23-31, May-June 2009.

[15] A. Shehu and A. Maraj, "Analysis of Different Parameters that Affect QoS in Different IPTV Architectures," Journal of Communication and Computer, vol. 8, pp. 1001-1012, 2011.

[16] T. Silverston, O. Fourmaux, A. Botta, A. Dainotti, A. Pescapé, G. Ventre and K. Salamatian, "Traffic Analysis of Peer-to-Peer IPTV Communities," Computer Networks, vol. 53, no. 4, pp. 470-484, 2009.

[17] A. Bikfalvi, J. Reinoso, I. Vidal, F. Valera and A. Azcorra, "P2P vs. IP Multicast: Comparing Approaches to IPTV Streaming Based on TV Channel Popularity," Computer Networks, vol. 55, no. 6, pp. 1310-1325, 2011.

[18] M. Cha, P. Rodriguez, J. Crowcroft, S. Moon and X. Amatriain, "Watching Television Over an IP Network," Proceedings of the 8th ACM SIGCOMM Conference on Internet Measurement, New York, USA, pp. 71–84, 2008.

[19] Y. Lu, F. Kuipers, M. Janic and P. Mieghem, "E2E Blocking Probability of IPTV and P2PTV," NETWORKING 2008, LNCS 4982, pp. 445-456, 2008.

[20] Y.-Y. Chen, "A Design of IPTV Conditional Access Mechanism Based on P2P Network," Peer-to-Peer Networking and Applications, vol. 7, no. 1, pp. 24-30, 2012.

[21] Y. Chen, E. Merrer, Z. Li, Y. Liu and G. Simon, "OAZE: A Network-friendly Distributed Zapping System for Peer-to-Peer IPTV," Computer Networks, vol. 56, no. 1, pp. 365-377, 2012.

[22] X. Hei, Y. Liu and K Ross, "IPTV over P2P Streaming Networks: The Mesh-Pull Approach," IEEE Communications Magazine, vol. 46, no. 2, pp. 86-92, February 2008.

[23] Q. Qi, Y. Cao, T. Li, X. Zhu and J. Wang, "Soft Handover Mechanism Based on RTP Parallel Transmission for Mobile IPTV Services," IEEE Transactions on Consumer Electronics, vol. 56, no. 4, pp. 2276- 2281, 2010.

[24] A. Bikfalvi, J. García-Reinoso, I. Vidal and F. Valera, "A Peer-to-Peer IPTV Service Architecture for the IP Multimedia Subsystem," International Journal of Communication Systems, vol. 23, no. 6–7, pp. 780-801, 2010.

[25] M. Cha, P. Rodriguez, S. Moon and J. Crowcroft, "On Next-Generation Telco-Managed P2P TV Architectures," Proceedings of the 7th International Conference on Peer-to-Peer Systems (IPTPS'08), pp. 1-6, 2008.

[26] C. Cho, I. Han, Y. Jun and H. Lee, "Improvement of Channel Zapping Time in IPTV Services Using the Adjacent Groups Join-Leave Method," International Conference on Advanced Communication Technology, vol. 2, pp. 971–975, 2004.

[27] J. M. Boyce and A. M. Tourapis, "Fast Efficient Channel Change [set-top box applications]," International Conference on Computers, Education Digest of Technical Papers, pp. 1–2. Jan. 2005.

[28] Hyunchul Joo, Hwangjun Song, Dai-Boong Lee and Inkyu Lee, "An Effective IPTV Channel Control Algorithm Considering Channel Zapping Time and Network Utilization," IEEE Transactions on Broadcasting, vol. 54, no. 2, pp. 208-216, June 2008.

[29] X. Hei, C. Liang, J. Liang, Y. Liu and K. W. Ross. "A measurement Study of a Large-scale P2P IPTV System," IEEE Transactions on Multimedia, vol. 9, no. 8, pp. 1672–1687, Dec. 2007.

[30] Z.Wang, Y. Huo, P. Liu, Z. Fan and Y. Ma, "Measurement-based Peer Selection for P2P-IPTV Services in Campus Networks," The Journal of China Universities of Posts and Telecommunications, vol. 18, Suppl. 2, pp. 61-65, 2011.

[31] L. Vu, I. Gupta, K. Nahrstedt and J. Liang, "Understanding Overlay Characteristics of a Large-Scale Peer-to-Peer IPTV System," ACM Transactions on Multimedia Computing, Communications and Applications, vol. 6, no. 4, pp. 31:1-31:24, 2010.

[32] X. Tu, H. Jin, J. Cao, S. Guo, L. Zheng and Z. Lv, "An Efficient Data Scheduling Scheme for P2P Storage-Constrained IPTV System," IEEE Transactions on Systems, Man and Cybernetics: Systems. vol. 43, no. 2, pp. 379-389, March 2013.

[33] Baochun Li, Zhi Wang, Jiangchuan Liu and Wenwu Zhu, "Two Decades of Internet Video Streaming: A Retrospective View," ACM Transactions on Multimedia Computing, Communications, and Applications (TOMCCAP) vol. 9, no. 1s, pp. 33:1-33:20, 2013.

[34] B. Song, M. Hassan and E.-N. Huh, "Delivering IPTV Service over a Virtual Network: A Study on Virtual Network Topology," Journal of Communications and Networks, vol. 14, no. 3, pp. 319-335, 2012.

[35] Fati, Suliman Mohamed, Rahmat Budiartu and Putra Sumari, "Provisioning Virtual IPTV Delivery Networks Using Hybrid Genetic Algorithm," in Proceedings of the 8th International Conference on Ubiquitous Information Management and Communication, p. 106, ACM, 2014.

[36] Hao Yin, Xuening Liu, Tongyu Zhan, Vyas Sekar, Feng Qiu, Chuang Lin, Hui Zhang and Bo Li, "Design and Deployment of a Hybrid CDN-P2P System for Live Video Streaming: Experiences with LiveSky," Proceedings of the 17th ACM International Conference on Multimedia, pp. 25-34, Beijing, China, 2009.

[37] T. Guo, J. Cai, M. Gao and C. H. Foh, "Dynamic Scheduling of a Mixture of Scalable IPTV and VoD Traffic over Wireless Home Networks," Proceedings of the 10th International Conference on Mobile and Ubiquitous Multimedia, pp. 123-126, 2011.

[38] Frank, Benjamin, Ingmar Poese, Yin Lin, Georgios Smaragdakis, Anja Feldmann, Bruce Maggs, Jannis Rake, Steve Uhlig and Rick Weber, "Pushing CDN-ISP Collaboration to the Limit," ACM SIGCOMM CCR, vol. 43, no. 3, 2013.

[39] Balachandran, Athula, Vyas Sekar, Aditya Akella and Srinivasan Seshan, "Analyzing the Potential Benefits of CDN Augmentation Strategies for Internet Video Workloads," in Proceedings of the Conference on Internet Measurement , pp. 43-56, ACM, 2013.

[40] Fernando M. V. Ramos, "GREEN IPTV: A Resource and Energy Efficient Network for IPTV," Technical report, University of Cambridge, 2013.

[41] Tsilopoulos, Christos, George Xylomenos and George C. Polyzos, "Are Information-Centric Networks Video-Ready?," Proc. of the IEEE 20th International Packet Video Workshop (PV), pp. 1-8, 2013.

[42] M. Jeffrey, S. Park, K. Lee, G. Adams and S. Savage, "Content Security for IPTV," IEEE Communications Magazine, vol. 46, no. 11, pp. 138–146, 2008.

[43] J. Park, "Subscriber Authentication Technology of AAA Mechanism for Mobile IPTV Service Offer," Telecommunication Systems, vol. 45, no. 1, pp. 37–45, 2010.

[44] http://www.dvb.org/, Accessed on 24 January 2016.

[45] M. Chung, Y. Lee, T. Shon and J. Moon, "A Security Model for IPTV with One-time Password and Conditional Access System for Smart Mobile Platform," Telecommunication Systems, vol. 52, no. 2, pp. 1399 – 1407, 2013.

[46] M. Jeffrey, S. Park, K. Lee, G. Adams and S. Savage, "Content Security for IPTV," IEEE Communications Magazine, vol. 46, no. 11, pp. 138-146, November 2008.

[47] R. Pippal and S. Tapaswi, "Secure Key Exchange Scheme for IPTV Broadcasting," Informatica, vol. 36, pp. 47–52, 2012.

[48] A. Pinto and M. Ricardo, "Secure Multi-cast in IPTV Services," Computer Networks, vol. 54, no. 10, pp. 1531–1542, 2010.

[49] A. Pinto and M. Ricardo, "On Performance of Group Key Distribution Techniques when Applied to IPTV Services," Computer Communications, vol. 34, no. 14, pp. 1708–1721, 2011.

[50] S.-C. Kim, S.-S. Yeo and S. Kim, "A Hybrid User Authentication Protocol for Mobile IPTV Service," Multimedia Tools and Applications, vol. 2, no. 2, pp. 283 – 296, July 2013.

[51] http://www.atsc.org/, Accessed 24 Jan. 2016.

[52] International Telecommunication Union, IPTV Focus Group Proceedings, T-PROC-IPTVFG-2008-PDF-E, 2008.

[53] A. Elmisery and D. Botvich, "Multi-agent Based Middleware for Protecting Privacy in IPTV Content Recommender Services," Multimedia Tools and Applications, vol. 64, no. 2, pp. 249-275, May 2013.

[54] Lee, Hyeungwoo, Jae-Young Choi and Jongpil Jeong, "On Cost-Reduced Channel Changing for Mobile IPTV Services in LTE-Advanced Systems," Proc. of the 7th International Conference on Advances in Future Internet (AFIN 2015), pp. 66-71, 2015.

[55] L. X. Cai, L. Cai, X. Shen and J. Mark, "Resource Management and QoS Provisioning for IPTV Over mmWave-based WPANs with Directional Antenna," Mobile Networks and Applications, vol. 14, no. 2, pp. 210–219, 2009.

[56] O. Karimi and J. Liu, "Power Efficient High Quality Multi-media Multi-cast in LTE Wireless Networks," Proc. of the 8th IEEE International Conference on Mobile Ad-Hoc and Sensor Systems, pp. 161-163, 2011.

[57] http://www.ericsson.com/spotlight/media/blog/opportunities-video-over-lte-enable-tv-anywhere-ericsson/, Accessed 24 Jan. 2016.

[58] http://www.iptv-news.com/2013/02/mwc-ericsson-debuts-broadcast-video-over-lte/, Accessed 24 Jan. 2016.

[59] http://www.fiercecable.com/story/verizons-mcadam-we-will-support-linear-tv-market-moving-ott/2015-09-17, Accessed 24 Jan. 2016.

[60] http://www.ericsson.com/news/1680666, Accessed 24 Jan. 2016.

[61] https://help.netflix.com/en/node/306, Accessed 24 Jan. 2016.

[62] J. Huang, F. Qian, Y. Guo, Y. Zhou, Q. Xu, Z. M. Mao, S. Sen and O. Spatscheck, "An In-depth Study of LTE: Effect of Network Protocol and Application Behavior on Performance," SIGCOMM'13, Hong Kong, China, pp. 363-374, 12–16 August 2013.

[63] R. Radhakrishna and A. Nayak, "An Efficient Video Adaptation Scheme for SVC Transport over LTE Networks," IEEE Int. Conf. on Parallel and Distributed Systems, pp. 127-133, Dec. 2011.

[64] R. Radhakrishna, B. Tirouvengadam and A. Nayak, "Channel Quality-based AMC and Smart Scheduling Scheme for SVC Video Transmission in LTE MBSFN Networks," IEEE International Conference on Digital Communications (ICC), pp. 6514-6518, 2012.

[65] S. Hwang, "Content and Service Protection for IPTV," IEEE Transactions on Broadcasting, vol. 55, no. 2, pp. 425-436, June 2009.

[66] UUSEE Inc. http://www.uusee.com, Accessed 24 Jan. 2016.

[67] C. Wu, B. Li and S. Zhao, "Exploring Large-scale Peer-to-Peer Live Streaming Topologies," ACM Transactions on Multi-media Computing, Communications and Applications (TOMM), vol. 4, no. 3, pp. 19:1-19:23, 2008.

APPENDIX A: ACRONYMS

Table 2. List of acronyms used in this paper.

Acronym	Description	Acronym	Description
AP	Access Point	NGN	Next-Generation Network
AVC	Advanced Video Coding	P2P	Peer to Peer
CABAC	Context-Adaptive Binary Arithmetic Coding	PC	Personal Computer
CN	Core Network	PEP	Performance Enhancing Proxy
CQI	Channel Quality Indicator	PTM	Point-To-Multipoint
eMBMS	evolved Multi-media Broadcast and Multi-cast Services	PTP	Point-To-Point
eNB	evolved Node B	QoE	Quality of Experience
GBR	Guaranteed Bit Rate	QoS	Quality of Service
HDTV	High Definition TeleVision	RAN	Radio Access Network
IDR	Instantaneous Decoding Refresh	SHE	Super Headend
IMS	IP Multi-media Subsystem	SNR	Signal-to-Noise Ratio
IPTV	Internet Protocol TeleVision	STB	Set Top Box
ISP	Internet Service Provider	SVC	Scalable Video Coding
LTE	Long Term Evolution	UE	User Equipment
MBS	Multi-cast and Broadcast Services	VoD	Video on Demand
MPEG	Moving Pictures Expert Group	VSO	Video Serving Office
NAT	Network Address Translation	WiMAX	World interoperability for Microwave Access

ملخص البحث:

بعد النجاح الذي أحرز في نقل الصوت عبر شبكات بروتوكولات الإنترنت، تحوّل حلم التلفاز القائم على بروتوكولات الإنترنت إلى واقع عبر جهود البحث المكثفة. واليوم، تبذل مؤسسات التقييس والباحثون والشركات التجارية جهوداً ضخمة لإنجاح التلفزة المبنية على بروتوكولات الإنترنت. وترجع النجاحات التي تحققت في هذا الميدان إلى العديد من العوامل، التي تتضمن بصورة رئيسية التقنيات المتقدمة في مجال تشفير الصور، والازدياد المستمر في عرض نطاق الإنترنت الذي يمكن للمستخدمين النهائيين الاستفادة منه، إلى جانب البروتوكولات والمعماريات الناضجة المتعلقة بالتشبيك السلكي واللاسلكي. وعلى الرغم من أن التلفزة المستندة على بروتوكولات الإنترنت لا تزال تواجه العديد من التحديات، فقد بدأ العديد من مزوّدي المحتوى ومزوّدي خدمات الإنترنت بتقديم خدمات التلفزة القائمة على بروتوكولات الإنترنت للزبائن.

إلا أن الطريق لا يزال طويلاً والحاجة ماسة إلى المزيد من الجهود. تبحث هذه الورقة في مدى واسع من مجالات البحث المتعلقة بالتلفزة القائمة على بروتوكولات الإنترنت. فهي تناقش أولاً البنية الخدمية العامة لخدمات التلفزة القائمة على بروتوكولات الإنترنت و المسائل التشخيصية المرتبطة بها. ثم تحوّل الانتباه إلى مسح تقنيات التشفير التي يمكن أن تكون أساساً للتلفزة المبنية على بروتوكولات الإنترنت في المستقبل، وبخاصة تشفير الصور القابل للتدريج. بعد ذلك، تناقش الورقة طرق توزيع مختلفة للتلفزة المبنية على بروتوكولات الإنترنت، بما في ذلك شبكات التوزيع من نوع نظير إلى نظير و شبكات توزيع المحتوى. ثم تتناول الورقة العمل الرئيسي في مجال المسائل المرتبطة بالأمان والخصوصية، وتبحث أخيراً في مجال الإنجازات الرئيسية في ميدان خدمات التلفزة المستندة إلى بروتوكولات الإنترنت في الشبكات اللاسلكية، مع التركيز على تقديم الخدمات المبنية على التطور في المدى الطويل.

ونعتقد أن مثل هذا المسح سيكون ذا فائدة كبيرة للباحثين الذين لديهم الرغبة في تثقيف أنفسهم بصورة إجمالية في مجال التلفزة القائمة على بروتوكولات الإنترنت قبل أن يتعمقوا في معالجة المشكلات البحثية المفتوحة؛ من أجل تقريب التلفزة المبنية على بروتوكولات الإنترنت أكثر فأكثر من الواقع.

Jordanian Journal of Computers and Information Technology (JJCIT), Vol. 2, No. 2, August 2016.

ENHANCEMENT OF TRIANGLE COORDINATES FOR TRIANGLE FEATURES FOR BETTER CLASSIFICATION

Nur Atikah Arbain[1], Mohd Sanusi Azmi[2], Laith Bany Melhem[3], Azah Kamilah Muda[4] and Hasan Rashaideh[5]

[1,2,3,4] Faculty of Information Technology and Communication,
Universiti Teknikal Malaysia Melaka, Malaysia.

[5] Al-Balqa Applied University, Al-Salt, Jordan.
nuratikah9.arbain@gmail.com[1], sanusi@utem.edu.my[2], lnm1989@gmail.com[3], azah@utem.edu.my[4], rashaideh@bau.edu.jo[5]

(Received: 25-Nov.-2015, Revised: 14-Jan.-2016, Accepted: 27-Jan.-2016)

ABSTRACT

Recently, the triangle features have been applied in digit recognition by adopting the angle as a part of the features. Most of the studies in digit recognition area which applied these features have given impressive results. However, the issue of big gap values that occurred between angles, ratios and gradients has shown a strong impact on the accuracy of the results. Therefore, we introduce our proposed method which is data normalization that has adopted the nature of triangle geometry in order to resolve this issue. Besides, we have applied other techniques, such as Z-score, Minimax and LibSVM function in the experiment. There are four digit datasets used which are HODA, MNIST, IFHCDB and BANGLA. The results of classification have shown that our proposed method has given better results compared to other techniques.

KEYWORDS

Triangle features, Triangle geometry, Feature extraction, Feature normalization, Feature scaling.

1. INTRODUCTION

1.1 Feature Geometry

Feature geometry is a compound of the basic geometric building blocks (e.g., points, lines, curves, surfaces and polygons) [1]. It was developed by [2], where natural classes are represented by hierarchical structures as well as by features themselves which have represented a major revision of the theory proposed by [3].

Feature geometry is an assortment of points that may produce a kind of multipoint, polyline or polygon geometry. Points are defined and created using the x, y coordinates, M is the measured value and Z is the vertical location. Every geometry kind has a definition that helps verify its validity. The tool should be designed to consider the geometry types. Whether or not it changes the geometry doesn't matter, however it has to handle the choice.

1.2 Properties of Triangle

A triangle is a closed figure that consists of three points which the three line segments have linked end-to-end. It is also known as a three sided polygon. There are eight properties or features for a triangle, which are: vertex, base, altitude, median, area, perimeter, interior angles and exterior angles. Generally, there are three types of triangle, which are: equilateral triangle, isosceles triangle and scalene triangle [4]. Nevertheless, there are three other types of triangle which can be formed from a scalene triangle. These are: acute scalene triangle, obtuse scalene triangle and isosceles right triangle. Table 1 shows the description of types of scalene triangle, while Figure 1 shows examples of types of triangle.

Table 1. Description of types of scalene triangle.

Types of Scalene Triangle	**Description**
Acute Scalene Triangle	This triangle has three angles each of which is less than 90 degrees.
Obtuse Scalene Triangle	This triangle has an angle which is more than 90 degrees.
Isosceles Right Triangle	This triangle has one right angle (90°).

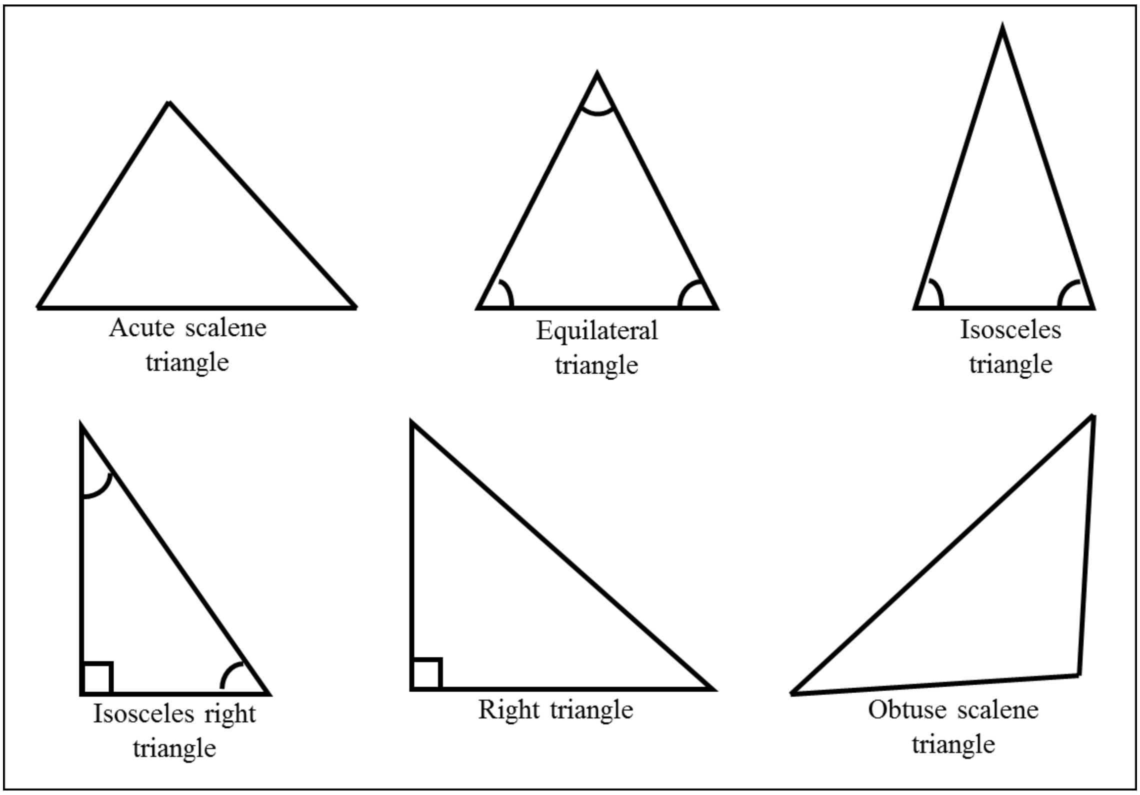


Figure 1. Examples of types of triangle.

Briefly, the vertex is known as a corner. It also refers to a point where lines meet. Usually, the vertex is used to mark the corners of a polygon. In addition, the included angle at each vertex is known as an interior angle of the polygon. Figure 2 shows an example of included angle [5].

However, in this paper, we are adopting scalene triangle, because the digit form is not symmetrical on the left and right sides like our eyes and nose, as stated in [6]. In with the [7], the author has reported that the Arabic digit and calligraphy do not have any pattern with the potential to form three edges for a triangle. Moreover, the author [7] has used nine features which will be described in detail in the feature extraction topic.

Jordanian Journal of Computers and Information Technology (JJCIT), Vol. 2, No. 2, August 2016.

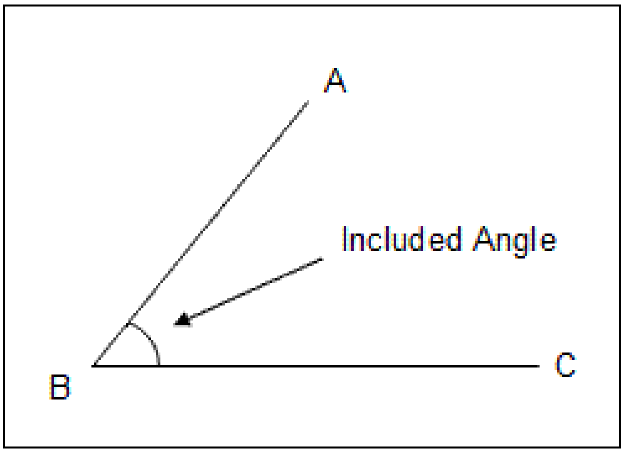


Figure 2. Included angle [5].

1.3 Feature Extraction

Feature extraction is an important task in image processing, because of the meaningful features extracted are vital in representing an object. An object is modelled and represented by geometrical properties. Geometry has properties that can be used in object recognition. The triangle properties have been adopted by many researchers to generate the proposed features for image classification. The triangle geometry is widely used in biometric research, such as face and fingerprint recognition [8]–[12]. Apart from that, other researchers have also adopted the similar geometry for intrusion [13], vehicle detection [14] and digit recognition [16]–[17].

In this paper, the triangle geometry features are used to extract the features from digit datasets. The triangle features have been proposed by [18]-[19]. These nine features of triangle are shown in Table 2.

The features in Table 2 will be applied to the several zones in order to increase the accuracy of classification by using Support Vector Machine (SVM). In this case, the Cartesian Plane Zone has been chosen as the initial method in conducting the experiments. It also can be applied to other zoning methods. Therefore, this paper is focusing on how the triangle features are normalized or being scaled. In this paper, we also explicate the results of classification after using several techniques as aforementioned.

The triangle geometry used by [17]-[18] is based on the Scalene Triangle method. In [18], 21 features have been used based on the Scalene Triangle method. However, only three out of 21 features are directly used in the experiments. The features are: angles of corners which have been labeled as A, B and C, ratios of sides $(a \times b,\ b \times c,\ c \times a)$ and gradient of side for each angle added by authors in [17].

The feature values of angles A, B and C are big compared to the features in 1, 2, 3, 7, 8 and 9. This can be proven based on the example of results in Table 3. The huge gap between angles of corners, ratios of sides and gradients proposed by [18] has imposed a big impact on the accuracy of classification. Thus, the features need to be scaled in order to improve the accuracy of result. By using HODA train dataset which consists of tr_10009_0.bmp, tr_10013_0.bmp, tr_10014_0.bmp and tr_10022_0.bmp, the samples of extracted features that used the triangle features in [17] are illustrated in Table 3.

In Table 3, the values of $\angle$A, $\angle$B and $\angle$C have shown that both angle gaps between ratios and gradients are very obvious. Due to this issue, the angles must be scaled in order to lessen the gaps. Recently, there are many normalization algorithms which can be applied to extract the features; for example, Z-score, libSVM, etc. However, this paper has reported the normalization technique by using the original extracted values which has been obtained from triangle geometry features, Z-score [20], libSVM scale function [21] and our proposed method.

Table 2. Triangle features [20].

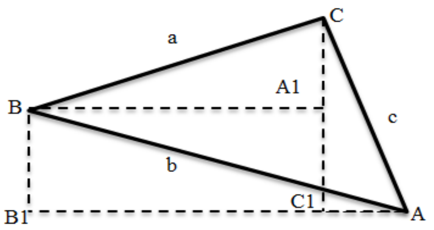
Triangle	No.	Feature	Formula
 $a = \sqrt{((A1(y) - C(y))^2) + ((A1(x) - B1(x))^2)}$ $b = \sqrt{((B1(y) - B(y))^2) + ((A(x) - B1(x))^2)}$ $c = \sqrt{((C1(y) - C(y))^2) + ((A(x) - C1(x))^2)}$	1	c:a	c:a = *c/a*
	2	a:b	a:b = *a/b*
	3	b:c	b:c = *b/c*
	4	A	$A = \arccos\frac{b^2+c^2-a^2}{2bc}$
	5	B	$B = \arccos\frac{a^2+c^2-b^2}{2ac}$
	6	C	$C = \arccos\frac{a^2+b^2-c^2}{2ab}$
	7	ΔBA	$\Delta BA = \frac{B(y)-C(y)}{B(x)-C(x)}$
	8	ΔBC	$\Delta BC = \frac{B(y)-A(y)}{B(x)-A(x)}$
	9	ΔCA	$\Delta CA = \frac{B(y)-C(y)}{B(x)-C(x)}$

Table 3. Example results of HODA train dataset.

c:a	a:b	b:c
0.63	0.62	2.55
1.05	0.49	1.92
2.06	0.89	0.54
0.97	0.51	2.02
0.63	0.62	2.55

i. Feature ratios

ΔBA	ΔBC	ΔCA
-0.67	-0.5	-1
-0.17	0	-0.33
0.25	0	0.5
-0.13	-0.25	0
-0.67	-0.5	-1

ii. Feature gradients

∠A	∠B	∠C
11.31	161.56	7.13
8.973	161.57	9.46
12.53	14.04	153.44
7.13	165.96	6.91
11.31	161.56	7.13

iii. Feature angles

The original features have been extracted by using triangle features shown in Table 3. These original features will be classified by using Support Vector Machine (SVM) without applying the normalization algorithm on the features. However, the triangle features will be normalized by using Z-score algorithm based on Equation (1). The normalized triangle features are worked out based on a linear scale and features will be scaled between ranges -1 to 1.

$$z = \frac{x - \mu}{\sigma} \quad (1)$$

2. Data Collection and Method

In this paper, we have used four different datasets that are taken from Arabic, Roman and Bangla handwritings. These datasets are known as HODA [22], Isolated Farsi Handwritten Character Database (IFHCDB) [23], MNIST [24] and BANGLA [25]. HODA and IFHCDB datasets are taken from Arabic handwriting. The MNIST dataset is taken from a Roman handwriting, while BANGLA dataset is taken from a Bangla handwriting. These datasets consist of alphabets and digits. However, in this paper, we are only focusing on digits instead of alphabets. The example of digits used in this paper is shown in Figure 3. Further details about these datasets can be referred to in [26].

2.1 Pre-processing

In the pre-processing, all datasets will be converted into a binary image by using Otsu's method [27]. After that, the datasets will get through the labelling process. The labelling process is a process to rename the images based on their type (test and train). In binary form, the foreground image is known as '1', while the background image is known as '0'. However, the binary images in HODA are different, whereas the foreground is white while the background is black. Thus, the images in HODA dataset will be inverted and the labelling process will take place thereafter. These datasets which are involved in pre-processing are shown in Figure 3.

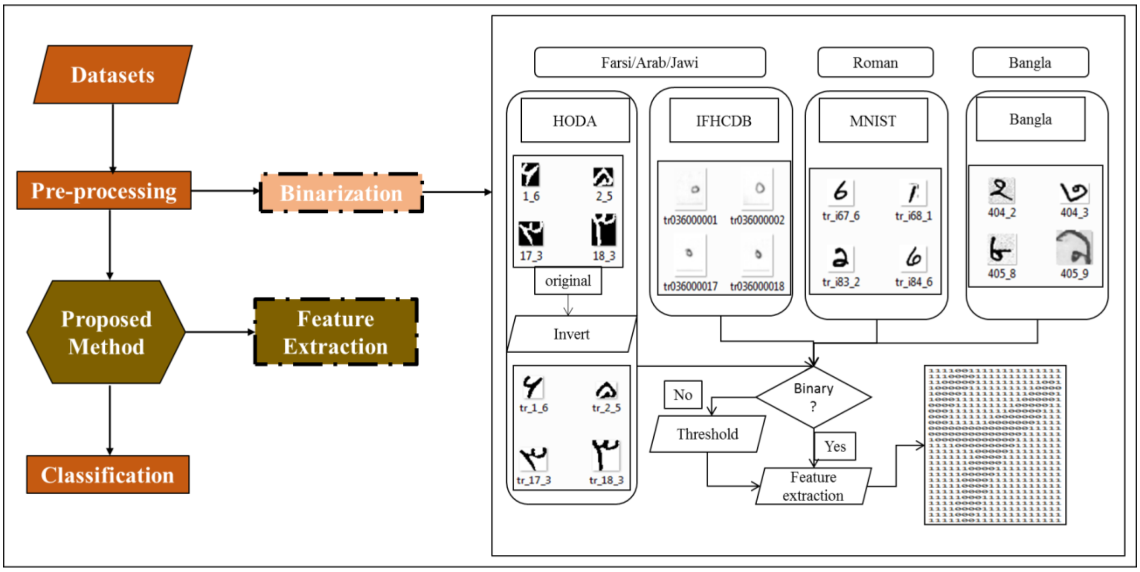


Figure 3. Pre-processing.

2.2 Proposed Method

In this paper, we have proposed a method based on the problem in [17]. The problem of huge gaps between the values of gradients, ratios of sides and angles of corners has imposed a strong impact on the accuracy of the results. However, we have only focused on the problem related to the angles of corners. As known, the total of a triangle's angles is 180 degrees. The total of a triangle's angles can be referred to in Equation (2).

$$\angle_{triangle} = \sum \angle A + \angle B + \angle C = 180 \quad (2)$$

In our proposed method, we have divided each angle by 180 degrees based on the total of a triangle's angles. This can be represented as shown in Equations (3), (4) and (5).

$$\angle A' = \angle A/180 \quad (3)$$

$$\angle B' = \angle B/180 \quad (4)$$

$$\angle C' = \angle C/180 \quad (5)$$

After diving each angle by 180 degrees, the triangle features represented in [18] have been changed. The triangle features in [17] are shown in Equation (6) while the new triangle features proposed are shown in Equation (7).

$$\mathrm{F}f = A_f, B_f, C_f, Rca_f, Rab_f, Rbc_f, \Delta\mathrm{BA}_f, \Delta\mathrm{BC}_f, \Delta\mathrm{CA}_f \tag{6}$$

$$\mathrm{F}f' = A_f', B_f', C_f', Rca_f, Rab_f, Rbc_f, \Delta\mathrm{BA}_f, \Delta\mathrm{BC}_f, \Delta\mathrm{CA}_f \tag{7}$$

In this paper, we have applied the Cartesian Plane Zone [28] as shown in Table 4. This zone has one main triangle and four zones represented as Zone A, Zone B, Zone C and Zone D.

Table 4. Cartesian plane zone.

Main Triangle				
111111111100111111111		11111111110	00111111111	
111111111000011111111		11111111100	00011111111	
111111110000001111111	**Zone B**	11111111000	00001111111	**Zone A**
111111110000000111111		11111111000	00x00111111	
111111000000000111111	Point B : (4 , 8)	11111100000	00000111111	Point B : (12 , 3)
111110000000000011111	Point C : (6 , 6)	11111000x00	00000011111	Point C : (14 , 6)
111100000111100001111	Point A : (8 , 5)	111100x0011	1110x001111	Point A : (16 , 8)
111100001111110001111		11110000111	11110001111	
111000011111111000111		1110x001111	111110x0111	
110000111111111000011		11000011111	11111000011	
100000111111111000001		10000011111	11111000001	
100001x111x1111100001		10000111111	11111100001	
00001111111001x110000		10000111111	11111100001	
000011111100001110000		00001111111	10011110000	
000111111000001110000		00001111110	00001110000	
000111111000001110000	**Zone D**	00x11111100	00001110000	**Zone C**
000111110000000100001		00011x11100	00001x10x00	
101111000000000000001	Point B : (2 , 14)	00011111000	00x00100001	Point B : (12 , 16)
111110000000001000011	Point C : (5 , 15)	1011110x000	00000000001	Point C : (15 , 15)
111100000111111111111	Point A : (7 , 17)	11111000000	00001000011	Point A : (18 , 15)
		11110000011	11111111111	
Point B : (6 , 11)				
Point C : (10 , 11)				
Point A : (14 , 12)				

Based on Table 4, the triangle form can be performed based on the coordinates of triangle that are represented by points A, B and C in each zone and the main triangle. Figure 4 shows the triangle form based on the coordinates, while Table 5 shows the summary of formulation for each zone including the main triangle.

Based on Table 5, C_x and C_y are the coordinates of point C for angles x and y of the main triangle, while N_x and N_y are the width and height of the image that has been converted into binary form. The coordinate of C_x is used as a border for the horizontal plane, while the coordinate of C_y is used as a border for the vertical plane. Based on the combination of the four zones, there are 36 features produced. This has caused that the total of features for Cartesian Plane Zone is 45 features, including the main triangle. As information, these zones are formed based on the coordinates of point C which is known as the centroid of the zone. We can distinguish between them by a shaded image. The mark 'x' on the red line is point C which is the centroid for Zone A as shown in Figure 5. The mark 'x' at the right side is point A, whereas at the left side it is point B. As information, the coordinates of point C need to be identified first before the coordinates of point A and point B. The coordinate of point C is used as a divider between right and left sections. The centroid at the right section represents point A, while the centroid at the left section represents point B. These features will be extracted by using the triangle concept. The results of feature extraction have produced 45 features after applying the

Cartesian Plane Zone. These features have been applied to HODA, IFHCDB, MNIST and BANGLA datasets.

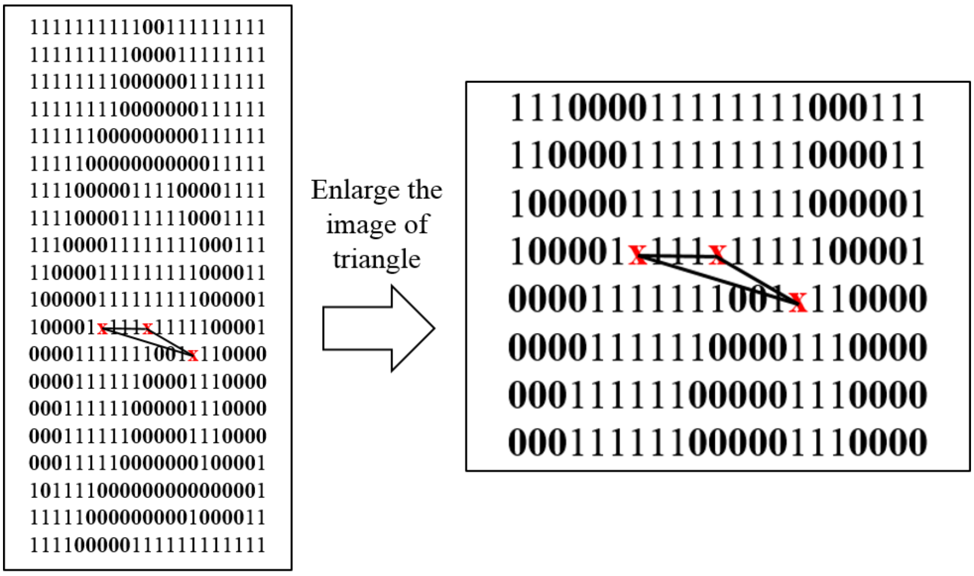


Figure 4. The form of triangle based on the coordinates.

Table 5. Summary of Cartesian plane zone.

Zone	Main Triangle	A	B	C	D
Binary Image	111111111100111111111 111111111000011111111 111111110000001111111 111111110000000111111 111111000000000111111 111110000000000011111 111100000111100001111 111100001111110001111 111000011111111000111 110000111111111000011 100000111111111000001 100001x111x1111100001 00001111111001x110000 000011111100001110000 000111111000001110000 000111111000001110000 000111110000000100001 101111000000000000001 111110000000001000011 111100000111111111111	00111111111 00011111111 00001111111 00x00111111 00000111111 00000011111 1110x001111 11110001111 111110x0111 11111000011 11111000001 11111100001	11111111110 11111111100 11111111000 11111111000 11111100000 11111000x00 111100x0011 11110000111 1110x001111 11000011111 10000011111 10000111111	11111100001 10011110000 00001110000 00001110000 00001x10x00 00x00100001 00000000001 00001000011 11111111111	10000111111 00001111111 00001111110 00x11111100 00011x11100 00011111000 1011110x000 11111000000 11110000011
Height (h)		$h = C_y$	$h = C_y$	$h = N_y - C_y + 1$	$h = N_y - C_y + 1$
Width (w)		$w = N_x - C_x + 1$	$w = C_x$	$w = C_x$	$w = N_x - C_x + 1$

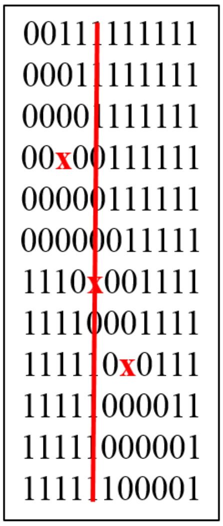


Figure 5. Zone A.

3. Experimental Setup and Results

In this paper, all the experiments were conducted by using Support Vector Machine (SVM). Table 6 shows the details of system requirements for SVM.

Table 6. System requirements for SVM.

Type	System Requirement
SVM Weka	Windows 7 and 8.1
	Ram size 4Gb and above
	Version 3.6.9
LibSVM tool (for grid search)	Windows 7 and 8.1
	Ram size 12Gb and above
	Python software (64bit, Version 3.5.0)
	Gnu plot: gp510-20150831 version
	Libsvm-3.17-GPU version, 64bit

The SVM classifier was chosen in order to compare the proposed method using the four datasets with the state-of-the-art works in this area. Cost and gamma are obtained from grid search using libSVM tool. There are three selected values of cost that were attained from grid search libSVM results, which are: c=32, c=8 and c=2, while gamma is g=0.00782. At this stage, the example of train and test images is provided by test datasets. The number of train and test images can be referred to in [26].

The first experiment has been conducted by using features without scaling, while the second experiment has been conducted by using Z-score algorithm. Table 7 shows the result without using normalization. It has been performed with an accuracy of not more than 70%. Among these datasets, HODA dataset has achieved the best result with an accuracy of 69.97%, while MNIST had the worst result with an accuracy of 52.06% in the first experiment. However, the result of classification shows some improvement that can be made by using either feature extraction or data normalization algorithm. Table 7 shows the results without using normalization, while Table 8 shows the results of normalization using Z-score.

Table 7 and Table 8 show that the triangle features without data normalization are slightly better compared to the results of data normalization by using Z-score algorithm. Thus, Z-score algorithm cannot be used as data normalization for triangle features. The third experiment has been conducted by using Minimax technique. We did not compare the results with those of

Jordanian Journal of Computers and Information Technology (JJCIT), Vol. 2, No. 2, August 2016.

SVM in [17]-[18], because the total of features used by [17]-[18] is not the same as used in this research. In [17]-[18], the researchers have used more than 200 features compared to only 45 features in this paper. Therefore, the results produced will affect the accuracy rate. Besides, this paper is introducing the data normalization technique that can be applied before the process of classification. Table 9 shows the results of normalization using Minimax.

Table 7. Results without normalization.

Dataset/Cost (C)	C=32	C=8	C=2
BANGLA	55.75	56.1	55.65
HODA	69.33	69.41	69.67
MNIST	52.06	52.14	52.34
IFHCDB	65.47	65.62	66.09

Table 8. Results of normalization using Z-Score.

Dataset/Cost (C)	C=32	C=8	C=2
BANGLA	55.7	56.03	55.85
HODA	69.34	69.41	69.65
MNIST	52.05	52.11	52.33
IFHCDB	65.4138	65.62	66.10

Table 9. Results of normalization using Minimax.

Dataset/Cost (C)	C=32	C=8	C=2
BANGLA	71.2	66.775	58
HODA	87.93	87.93	87.93
MNIST	74.02	71.86	69.23
IFHCDB	89.81	87.66	84.15

The fourth experiment has been conducted by using our proposed method. Table 10 shows that HODA and IFHCDB for Arabic digit datasets have achieved an accuracy of 90.35% and 91.72%, respectively compared to other datasets. All datasets which utilized our proposed normalization method outperformed other methods. Next, Table 11 shows the comparison results using a cost value of 32 and a gamma value of 0.078125, while Table 12 shows the training time taken to build each technique.

Table 10. Results of proposed method (normalization using LibSVM).

Dataset/Cost (C)	C=32	C=8	C=2
BANGLA	77.3	75.1	71.65
HODA	90.35	89.1	87.55
MNIST	77.91	76.42	73.86
IFHCDB	91.72	91.25	90.57

Table 11. Comparison results.

Dataset	Without normalization using LibSVM	Normalization using Z-Score	Normalization using Minimax	Normalization using LibSVM (Proposed Method)
BANGLA	55.75	55.7	71.2	77.3
HODA	69.33	69.34	87.925	90.35
MNIST	52.06	52.05	74.02	77.91
IFHCDB	65.47	65.41	89.81	91.72

Table 12. Training time taken for each technique (in seconds).

Dataset	Without normalization using LibSVM	Normalization using Z-Score	Normalization using Minimax	Normalization using LibSVM (Proposed Method)
BANGLA	33.16	277.36	508.69	343.1
HODA	103.68	6203.69	8154.28	6005.79
MNIST	492.07	8304.22	8747.97	7771.47
IFHCDB	7.24	233.94	333.79	138.7

Based on Table 12, the training time taken for our proposed method was remarkably longer compared to the case without normalization using LibSVM technique. However, the results without normalization were worse compared to other techniques. Besides, the time taken for our proposed method has shown to be shorter in general compared to normalization using Z-Score and Minimax techniques.

4. CONCLUSION

This paper is extended from our previous article [28]. Overall, we report that our normalization method using triangle features is suitable to our feature extraction method. Previously, we have published several papers on techniques that have been used in the extraction process. However, we have never mentioned any normalization techniques which are used to enhance the accuracy of classification for digit datasets. Thus, in this paper, we have stated that we have used normalization technique in our features before the process of classification. We used our proposed normalization algorithm in [18] without reporting it. Thus, any benchmarking should be referred to in paper [18]. In [18], we have introduced 25 zones in order to improve the accuracy of classification. Based on Table 11, the results have shown that our proposed method has better results compared to other techniques. Thus, it has been proven that our proposed method which is based on the nature of summation of angles A, B and C that are equal to 180 degrees has given better results using LibSVM technique. The next step will consider further improvements in different rotation invariants.

ACKNOWLEDGEMENTS

The authors would like to thank the Ministry of Education for funding this research (grant FRGS/2/2014/ICT/02/FTMK/02/F00246). Thanks are also extended to Zamalah Scheme Scholarship for providing the scholarship for this research. Besides, thanks are due to Universiti Teknikal Malaysia Melaka and the Faculty of Information Technology and Communication for providing excellent research facilities.

REFERENCES

[1] G. Gartner, L. Meng and M. P. Peterson, Digital Terrain Modelling, Lecture Notes in Geoinformation and Cartography, 2007.

[2] G. N. Clements, The Geometry of Phonological Features, Cambridge University Press, vol. 2, pp. 225–252, 1985.

[3] M. L. T. Cossio, L. F. Giesen, G. Araya, M. L. S. Pérez-Cotapos, R. L. VERGARA, M. Manca, R. A. Tohme, S. D. Holmberg, T. Bressmann, D. R. Lirio, J. S. Román, R. G. Solís, S. Thakur, S. N. Rao, E. L. Modelado, A. D. E. La, C. Durante, U. N. A. Tradición, M. En, E. L. Espejo, D. E. L. A. S. Fuentes, U. A. De Yucatán, C. M. Lenin, L. F. Cian, M. J. Douglas, L. Plata and F. Héritier, Uma ética para quantos?, vol. XXXIII, no. 2, pp. 81–87, 2012.

[4] E. W. Weisstein, "Triangle," Wolfram Research, Inc., 1999-2016.

[5] "Math Open Reference Citations," [Online], Available: http://www.mathopenref.com/site/cite.html, [Accessed: 19-Nov-2015].

[6] C. Lin and K. Fan, "Human Face Detection Using Geometric Triangle Relationship," Proc. 15th Int. Conf. Pattern Recognition (ICPR-2000), vol. 2, pp. 941–944, 2000.

[7] Moh'd Sanusi Azmi, Khairuddin Omar, M. F. Nasrudin and A. K. Muda, Fitur Baharu Dari Kombinasi Geometri Segitiga Dan Pengezonan Untuk Paleografi Jawi Digital, Ph.D. dissertation, Universiti Kebangsaan Malaysia, 2013.

[8] L. Gao and Y. Xu, "Face Orientation Recognition Based on Multiple Facial Feature Triangles," in Proc. of the 2012 Int. Conf. Control Eng. Commun. Technol. (ICCECT 2012), pp. 928–932, 2012.

[9] Z. Liu and J. An, "A New Algorithm of Global Feature Matching Based on Triangle Regions for Image Registration," pp. 1248–1251, 2010.

[10] X. Chen, J. Tian, X. Yang and Y. Zhang, "An Algorithm for Distorted Fingerprint Matching Based on Local Triangle Feature Set," IEEE Transactions on Information Forensics and Security, vol. 1, issue. 2, pp. 169–177, 2006.

[11] Z. Zhang, S. Wang and A. I. Morphing, "Multi-feature Facial Synthesis Based on Triangle Coordinate System," Proc. of the IEEE 2nd International Conference on Computer Science and Network Technology (ICCSNT 2012), no. 2, pp. 141–145, 2012.

[12] M. M. M. Tin and M. M. Sein, "Multi-Triangle Based Automatic Face Recognition System By Using 3D Geometric Face Feature," in IEEE International Instrumentation and Measurement Technology Conference (I2MTC 2009), Singapore, pp. 895–899, 5-7 May 2009.

[13] P. Tang, R. Jiang and M. Zhao, "Feature Selection and Design Olintrusion Detection System Based on k-means and Triangle Area Support Vector Machine," Proc. of the IEEE 2nd International Conference on Future Networks, 2010. (ICFN’10), pp. 144–148, 2010.

[14] A. Haselhoff and A. Kummert, "A Vehicle Detection System Based on Haar and Triangle Features," in Intelligent Vehicles Sysmposium, pp. 261–266, 2009.

[15] S. Mozaffari, K. Faez and V. Margner, Application of Fractal Theory for On-Line and Off-Line FarsiDigit Recognition, Chapter: Machine Learning and Data Mining in Pattern Recognition, Vol. 4571 of the series Lecture Notes in Computer Science, pp. 868–882, 2007.

[16] R. Ebrahimpour, A. Esmkhani and S. Faridi, "Farsi Handwritten Digit Recognition Based on Mixture of RBF Experts," IEICE Electron Express, vol. 7, no. 14, pp. 1014–1019, 2010.

[17] M. S. Azmi, K. Omar, M. F. Nasrudin, B. Idrus and K. Wan Mohd Ghazali, "Digit Recognition for Arabic/Jawi and Roman Using Features from Triangle Geometry," AIP Conf. Proc., vol. 1522, pp. 526–537, 2013.

[18] M. S. Azmi, K. Omar, C. Wan and S. Bahri, "Exploiting Features from Triangle Geometry for Digit Recognition," Proc. of the IEEE 2013 International Conference on Control, Decision and Information Technologies (CoDIT), Hammamet, Tunisia, pp. 876 - 880, 6-8 May 2013.

[19] Norm Matloff, From Algorithms to Z-Scores: Probabilistic and Statistical Modeling in Computer Science, University of California, p. 464, 2013.

[20] C.-C. Chang and C.-J. Lin, "LIBSVM: A Library for Support Vector Machines," ACM Transactions on Intelligent Systems and Technol., vol. 2, pp. 27:1–27:27, 2011.

[21] H. Khosravi and E. Kabir, "Introducing a Very Large Dataset of Handwritten Farsi Digits and a Study on Their Varieties," Pattern Recognition Letters, vol. 28, no. 10, pp. 1133–1141, 2007.

[22] S. Mozaffari and H. Soltanizadeh, "ICDAR 2009 Handwritten Farsi/Arabic Character Recognition Competition," Proc. of the IEEE 10th International Conference on Document Analaysis and Recognition (ICDAR), pp. 1413–1417, 2009.

[23] Y. LeCun and C. Cortes, MNIST Handwritten Digit Database, 1992.

[24] U. Bhattacharya and B. B. Chaudhuri, "Handwritten Numeral Databases of Indian Scripts and

Multistage Recognition of Mixed Numerals," IEEE Transactions on Pattern Analaysis and Machine Intelligence, vol. 31, issue 3, pp. 444–457, 2009.

[25] M. S. Azmi, K. Omar, M. F. Nasrudin and A. K. Muda, "Features Extraction of Arabic Calligraphy Using Extended Triangle Model for Digital Jawi Paleography Analysis," International Journal of Computer Information Systems and Industrial Management Applications, vol. 5, pp. 696–703, 2013.

[26] P. Smith, D. B. Reid, C. Environment, L. Palo, P. Alto and P. L. Smith, "A Threshold Selection Method from Gray-Level Histograms," IEEE Transactions on Systems, Man and Cybernetics, vol. 9, issue 1, pp. 62–66, Jan. 1979.

[27] M. S. Azmi, K. Omar, M. F. Nasrudin, A. Abdullah and K. Wan Mohd Ghazali, "Features Extraction of Arabic Calligraphy Using Extended Triangle Model for Digital Jawi Paleography Analysis," International Journal of Computer Information Systems and Industrial Management Applications, vol. 5, pp. 696-703, 2013.

[28] M. S. Azmi, N. A. Arbain, A. K. Muda, Z. Abal Abas and Z. Muslim, "Data Normalization for Triangle Features by Adapting Triangle Nature for Better Classification," 2015 IEEE Jordan Conference on Applied Electrical Engineering and Computing Technologies (AEECT), pp. 1–6, 3-5 Nov. 2015.

ملخص البحث:

تم حديثاً استخدام خصائص المثلث في تمييز الأرقام عن طريق اعتبار الزاوية جزءاً من الخصائص. وقد أعطت غالبية الدراسات في مجال تمييز الأرقام التي استخدمت تلك الخصائص نتائج مشجّعة. ومع ذلك، فإنّ مسألة الفجوة الكبيرة التي حدثت بين الزوايا ونسب الأضلاع والميلانات كان لها تأثير قويّ على دقة النتائج. ولهذا، نقترح طريقةً في هذه الدراسة هي عبارة عن تسوية البيانات باستخدام طبيعة هندسة المثلث من أجل حلّ هذه المسألة. إلى جانب ذلك، طبّقنا تقنياتٍ أُخرى، مثل (Z-Score)، و(Minimax) و(LibSVM)، في التجارب المتعلقة بهذه الدراسة. وقد تم استخدام أربع مجموعاتٍ من البيانات الرقمية هي: "HODA" و"MNIST" و"IFHCDB" و"BANGLA". وبينت نتائج التصنيف أن الطريقة المقترحة أعطت نتائج أفضل مقارنة بالتقنيات الاخرى.

Jordanian Journal of Computers and Information Technology (JJCIT), Vol. 2, No. 2, August 2016.

ON THE LOCATION-AWARE COOPERATIVE SPECTRUM SENSING IN URBAN ENVIRONMENT

Raouf Abozariba[1], Mohammad N. Patwary[2], Abdel-Hamid Soliman[3] and Mohamed Abdel-Maguid[4]

[1,2,3] Faculty of Computer, Engineering and Sciences,
Staffordshire University, Stoke-on-Trent, UK.

[4] Department of Science and Technology
University Campus Suffolk, Ipswich, UK.
{r.abozariba[1], m.n.patwary[2], a.soliman[3]}@staffs.ac.uk, m.abdel-maguid@ucs.ac.uk[4]

(Received: 16-Dec.-2015, Revised: 10-Feb.-2016, Accepted: 21-Feb.-2016)

ABSTRACT

Spectrum sensing is a key enabling technology for cognitive radio networks (CRNs). The main objective of spectrum sensing is to provide more spectrum access opportunities to cognitive radio users without interfering with the operations of the licensed network. Spectrum sensing decisions can lead to erroneous sensing with low performance due to fading, shadowing and other interferences caused by either terrain inconsistency or dense urban structure. In order to improve spectrum sensing decisions, in this paper a cooperative spectrum sensing scheme is proposed. The propagation conditions such as the variance and intensity of terrain and urban structure between two points with respect to signal propagation are taken into consideration. We have also derived the optimum fusion rule which accounts for location reliability of secondary users (SUs). The analytical results show that the proposed scheme slightly outperforms the conventional cooperative spectrum sensing approaches.

KEYWORDS

Cooperative spectrum sensing, Location-aware, Cognitive radio, Signal propagation, Urban environment.

1. INTRODUCTION

Today's wireless networks are characterized by a fixed spectrum assignment policy. As a result of increasing demands for wireless applications, there is a lack of frequency resources. In recent years, we have seen a significant interest in quantitative measurements of licensed and unlicensed spectrum utilization. Several research groups, companies and regulatory bodies have conducted studies of varying times and locations with the aim to capture the overall utilization rate of spectrum. These studies have given a significant amount of insight into spectrum use [1], [2]. Most of these studies have shown that a large amount of allocated spectrum are under-utilized and create what is called *spectrum holes*, resulting in a waste of valuable frequency resources [3]–[5]. *Spectrum holes* represent the potential opportunities for non-interfering use of spectrum and are considered as multi-dimensional regions within frequency, time and space. Consequently, high blockage probabilities are unavoidable for many users due to shortages of frequency resources caused by inefficient utilization. Cognitive radio (CR) technology is introduced in the literature to solve these ongoing spectrum inefficiency problems. The term cognitive radio was first introduced by Mitola in the 1990s to take advantage of the under-utilized scarce wireless spectrum [6]. CR is a key enabling technology for dynamic spectrum access, which provides higher bandwidth to mobile users via heterogeneous wireless architectures [7].

There are three main CR paradigms for sharing the spectrum: interweave, overlay and underlay. In interweave paradigm, cognitive users opportunistically exploit the primary radio spectrum only when the primary signal is detected to be idle. In overlay paradigm, cognitive users help maintain and/or improve primary users' (incumbent users) communication while utilizing some spectrum resources for their own communication needs. The underlay paradigm allows cognitive users to share the frequency bandwidth of the primary network only if the resultant interference power level at the primary receiver is below a given threshold.

CR is performed by a cycle which consists of three main stages: spectrum sensing, dynamic spectrum allocation and transmit power control, see Figure 1. Spectrum sensing is considered as one of the most challenging tasks in CR technology [8]-[9]. In dynamic spectrum allocation, channels are allocated to users based on spectrum availability. This allocation also depends on internal and/or external policies between cooperative networks. Transmit power control enables CR transmission to be controlled at the beginning of and during the transmission. This enables CRNs to serve more users and to lower the interference to the spectrum owners [10]. In spectrum sensing, the performance is usually measured by two key factors: probability of detection and probability of false alarm. The former is a probability that the detector correctly detects the signal when it is present in a given band. On the contrary, probability of false alarm is a probability that the detector incorrectly detects the presence of a signal though it is actually in temporal/permanent idle state. Probabilities are usually represented in a plot of the probability of detection versus the probability of false alarm, which is commonly referred to as radio operating characteristics (ROC). In this paper, these two factors will be the basis for determining the reliability of the proposed scheme and the results will be compared with the performance of the conventional hard combining scheme. The main contributions of this paper are as follows:

- We analyze the effect of the SUs' locations on spectrum sensing.
- We derive the fusion rule with consideration of SUs' locations within cooperative cognitive spectrum sensing.
- We propose a location-aware cooperative sensing scheme that combines the sensing results from multiple SUs. The sensing results are considered according to the reliability measured by the location information.

The remainder of the paper is organized as follows. Back- ground and motivation are presented in the next section. In Section 3, we define the system model and assumptions of the cooperative CR network that is used in our analysis. Section 4 gives a review of our proposed sensing method. Analytical results are discussed in Section 5. Finally, we make our concluding remarks in Section 6.

2. BACKGROUND AND MOTIVATION

2.1 Spectrum Sensing

Spectrum sensing is considered as one of the most challenging tasks in CR technology [8]-[9]. In the literature, various spectrum sensing methods and algorithms have been investigated, each having different operational requirements, advantages and disadvantages. The most common sensing methods are: feature detector, matched filtering and energy detection. Feature detector is performed using cyclic spectrum density function of the received signal or by matching general features of the received signal to the already known primary signal characteristics. If the structure of the signal source is known, optimal detection in stationary Gaussian noise is achieved by matched filtering method and coherent detection. This type of coherent detection may be a viable approach for early CR deployment, where the secondary system is limited to operate within a few systems such as Television (TV) and Digital Video Broadcasting (DVB). However, if more bands are being opened for opportunistic access, the implementation cost and complexity associated with this approach will increase [11]. A simpler alternative for the detection of a signal in Gaussian noise is to employ energy detection, which has drawn more attention in recent years, mainly due to its low complexity [12]. Energy detection determines the existence or absence of PUs by comparing the received energy at a CR to a pre-defined threshold. The performance of the energy detection increases monotonically with the quality of the received signal [13]. In [9], energy detection technique has been tested in an environment of low signal to noise ratio (SNR), while in [14] sequential energy detection was proposed to reduce sensing time. The authors in [15]-[16] studied the performance of energy detection under different channel constraints, such as additive white Gaussian noise (AWGN) and fading channels. Measuring the power of the received signal is the only requirement for energy detection, which then can be compared with a pre-defined threshold [17].

2.2 Cooperative Spectrum Sensing

The main challenge faced today by CR researchers is the ability to detect and utilize spectrum opportunities on a non-interference basis. Constructive and/or destructive interference can occur when signals travel along different paths to reach receivers, which causes attenuation and delay to the signal.

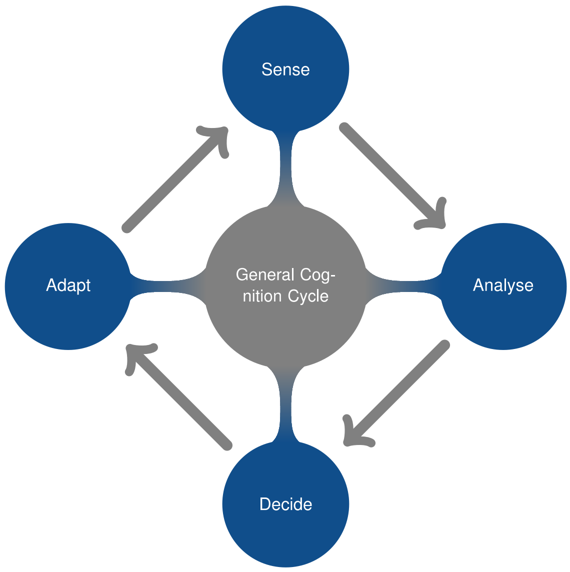


Figure 1. Cognitive cycle.

The received signal consists of several multi-path components, each of which is the result of the interaction of the transmitted waves with the surrounding environment. This issue has prompted researchers to turn to cooperative cognitive radio (CCR) networks, where all CRs collaborate in the spectrum sensing process. The advantage gained by using CCR networks lies at the achievable space diversity due to using multiple CRs [18]-[19]. In this context, cooperation indicates that a number of CRs are responsible to sense one particular channel at defined time and location. Cooperative sensing has gained interest in recent research papers, such as the work in [20]–[23]. Different cooperative sensing strategies have been studied to achieve better reliability of detecting primary signals. Sensing performance of a multiple primary user detector is discussed in [11]. Analytical formulae have been found for its false alarm probability and decision threshold. Numerical examples show significant performance gain over several detection algorithms in scenarios with realistic parameters. In [24], a weighted cooperative spectrum sensing scheme based on energy detection for minimizing the total error probability in CCR networks is investigated and analyzed, with resorting to allocation of optimal weight coefficients to individual cooperative secondary users.

Cooperative sensing is proposed in the literature as a solution to the problems that arise in spectrum sensing due to noise uncertainty, fading and shadowing [20]. However, the performance of cooperative spectrum sensing can be deceptive, because it highly depends on the reliability of the global decision. To address this challenge, various potential solutions were presented, as in [24] and [11]. In these studies, it is assumed that all secondary users are capable of estimating the received power with equal accuracy. However, such an assumption may not be always realistic, especially in high terrain and urban areas, where the structure of signal paths can vary dramatically. In this paper, we specifically address this issue and propose a new scheme to optimize spectrum sensing by considering location awareness. We show that the accuracy of spectrum sensing can be improved by avoiding secondary users' incorrect decisions caused by refraction and diffraction of primary signals. Furthermore, the proposed scheme takes advantage of spatial diversity raised due to the random distribution of secondary users within the coverage area.

Table 1. Notations used.

Notations	Descriptions
$P_{f,i}$	Probability of false alarm at the ith user
$P_{d,i}$	Probability of detection at the ith user
H_0	Null hypothesis
H_1	Alternative hypothesis
λ	Threshold
σ_w^2	Noise power
$Q(.)$	Q function
γ	Average signal to noise ratio
Y	Received energy for binary hypothesis
$r(t)$	Received energy
h	Channel gain
$s(t)$	Transmitted signal
$w(t)$	AWGN with zero mean and unit power
N	Sample Number
M	Number of SUs
u_0	Global binary decision at the SBS
u_i	Binary decision at the ith SU
Q_d	Overall detection probability
Q_f	Overall false alarm probability
S_j	Set of all decisions at the SBS which are equal to j where $j \in (0,1)$

3. System Model and Assumptions

In this paper, we consider an infrastructure-based CCR network which consists of one primary and one secondary network. A secondary base station (SBS) which also functions as a fusion center is also part of the secondary network. The network includes M number of secondary users (SUs), which are scattered in a given geographical area at the periphery of the coverage of the SBS.

Jordanian Journal of Computers and Information Technology (JJCIT), Vol. 2, No. 2, August 2016.

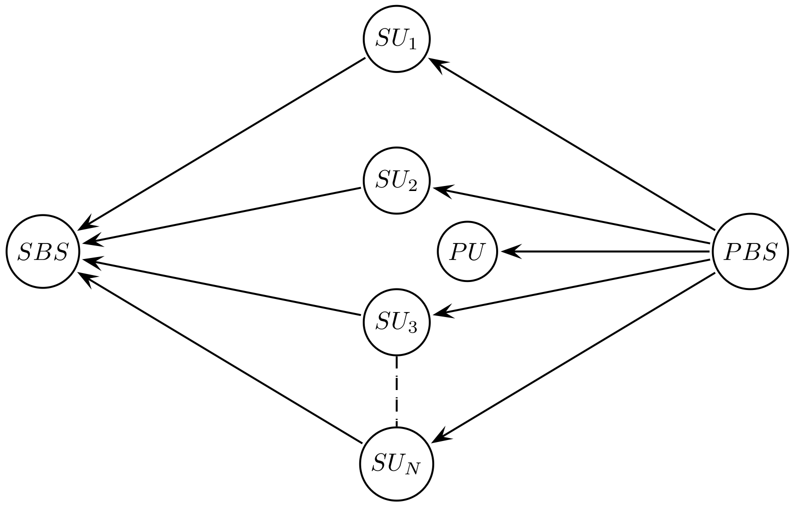


Figure 2. System model.

In Figure 2, SUs observe the same hypotheses independently and transmit their measurements to the SBS through a dedicated control channel which is assumed to maintain communication between SUs and their associated SBS. Here, the control channel is considered to be perfect. Similarly, the primary network consists of a primary base station (PBS) and primary users (PUs). Since we are interested in the downlink frequency channels of the primary network, SUs only perform spectrum sensing to target downlink channels (from the base station to the user), which are transmitted by the PBS. SBS decides whether a primary signal exists or not, which is a normal random process that depends on both the PBS activities and the spectrum sensing accuracy of SUs. Spectrum sensing at the SUs is performed using energy detection, which is commonly formulated as a Neyman-Pearson (NP) type binary hypothesis test problem. In such sensing technique, the received signal at each SU and at time *t* is given by [16].

$$r_i(t) = \begin{cases} w_i(t) & \text{if channel is free } H_0 \\ hs_i(t) + w_i(t) & \text{if channel is busy } H_1 \end{cases} \quad (1)$$

where h is the channel given, $s_i(t)$ is the transmitted signal and $w_i(t)$ is the AWGN with zero mean and unit power $\mathcal{N}(0,1)$. The hypothesis models H_0 and H_1 as presented in Equation (1) denote the absence and the presence of the primary signal, respectively. The performance measurement of any CR system is determined by its probability of detection $(P_{d,i})$ and probability of false alarm $(P_{f,i})$. High $(P_{d,i})$ guarantees minimal interference with primary network, and low $(P_{f,i})$ guarantees throughput improvement for the secondary network. Both measurements are used as the basis to determine the performance of CR systems in this paper. $P_{d,i}$ and $P_{f,i}$ can be estimated by:

$$P_{d,i} = \Pr\{Y > \lambda | H_1\} \quad (2)$$

and

$$P_{f,i} = \Pr\{Y > \lambda | H_0\}, \quad (3)$$

where *Y* is the received energy. The probability of detection in Equation (2) refers to the probability of accepting H_1 when H_1 is true. The probability of false alarm in Equation (3) refers to the probability of accepting H_1 when H_0 is true [9]. With direct computation of (2) and

(3), we have:

$$P_{d,i} = Q\left(\left(\frac{\lambda}{N{\sigma_w}^2} - 1 - \gamma\right)\frac{\sqrt{N}}{1+2\gamma}\right) \tag{4}$$

and

$$P_{f,i} = Q\left(\left(\frac{\lambda}{N{\sigma_w}^2} - 1\right)\sqrt{N}\right), \tag{5}$$

where λ is the threshold, ${\sigma_w}^2$ is the noise power, $Q(.)$ is the Q function, γ is the average signal to noise ratio and N is the sample number [16]. According to the information collected from SUs, the SBS makes its final decision about the spectrum availability. A specified decision method is adopted in order for the SBS to reach its final conclusion. Decision methods are generally divided into hard combination decision and soft combination decision. In hard combination decision, each SU reports its local decision to the SBS and the decision is made from a specific rule, such as logic "AND" and logic "OR". Hard combining is simple to implement and requires lower overhead (e.g., one-bit) [25]. For soft combination decision, the original observed data at the SUs, such as received power, is reported to the SBS and the decision is obtained by using one of the available techniques, such as equal gain combining (EGC) and log likelihood ratio (LLR) [25]–[27]. Soft combining method outperforms hard combining method in terms of the probability of missed opportunity. However, hard combining decisions are found to perform as good as soft decisions when the number of SUs is high [25]. In this paper, we consider hard combination decision as the core of our cooperative spectrum sensing decision method. In order to improve the accuracy of the chosen sensing method, we assume that the system is aware of the SU's location. SUs can be located in high dense built areas, where power measurements are less reliable due to various phenomena such as diffraction and reflection. It is important that the sensing decision method considers the SUs' locations to determine the environmental conditions of SUs, because the sensing accuracy is a function of location in respect to the source signal. Inaccurate sensing measurements, which are sent to the SBS, can potentially degrade the sensing accuracy. In a typical cellular network, the locations are stored in the HLR (Home Location Register). The HLR is the central user database in the mobile radio network. It stores the user and subscriber information. The location of both PBS and SUs can be described by longitude and latitude, which are a random collection of points on a coverage area [7]. The locations of PBSs can be obtained based on publicly available data, such as Consolidated Database System (CDBS). The locations of mobile SUs can be determined by various location estimation techniques, such as time-of-arrival (TOA), angle-of-arrival (AOA), received signal strength (RSS), pattern recognition and Bayesian filters [28].

4. LOCATION-AWARE COOPERATIVE SPECTRUM SENSING

4.1 Urban Propagation

Since spectrum is a very limited commodity in mobile communication systems, particularly in urban areas, we focus our study on urban environment [29]. Propagation of electromagnetic waves in urban areas in cellular frequencies is influenced by the geographical and structural area. Therefore, a detailed vector database of the buildings is required in order to establish a propagation map. Typically, the multi-path propagation is very important in urban environments. Urban propagation models already play an important role in the development, planning and deployment of mobile radio systems where coverage is the primary goal. Urban propagation models could also be used for signal detection reassessment, as we show in this paper. The attenuation of the signal strength in cellular frequencies is caused by three factors: path loss, multi- path fading and shadowing [30]. Here, we define the three attenuation factors:

- Path loss factor characterizes the rate at which the signal strength decays with the increase of the distance from a transmitter. Path loss factor increase is observed when signal propagation is subject to reflection and deflection from surrounding objects, such as floors, walls and trees.
- Multi-path fading, also called fast fading, is the propagation phenomenon that results in radio signals reaching the receiving antenna by two or more paths. This is caused by reception of multiple copies of a transmitted signal through multi-path propagation. An amplitude distribution is often described by a Rician or Rayleigh distribution, depending on whether a dominant component among the multiple copies exists or not. Usually, fast fading effect can be removed by averaging the received power over a time interval.
- Shadowing, often referred to as slow fading, represents a slow variation in a received signal strength due to obstacles in propagation paths. This factor increases the signal detection uncertainty.

4.2 Proposed Scheme

We propose a scheme, which is capable of improving the sensing accuracy of a CCR system. In this scheme, SUs determine their locations to realize the signal path quality in reference to the PBS (source signal). The location data of SUs are sent to the SBS for further investigation, see Figure 3. Knowledge of SUs' locations at the SBS can determine whether a line-of-sight (LOS) between transmitter and receiver exists and whether the path is obstructed by large building developments and structures such as wind turbines (e.g., Non- line-of-sight (NLOS) propagation), …etc., which can potentially cause the received signal to be less detectable at the SUs. Such consequential impact can degrade the sensing quality when considering a global decision.

4.3 Trust Value

In the proposed scheme, the sensing results from SUs are returned to the SBS along with location coordinates. We note that when SBS is in possession of the locations of SUs and PBS and PBS's networking information, including channel, height, transmit power antenna directionality …etc., the SBS will have the ability to approximate a trust value. There are a number of propagation models, which are well designed and give good accuracy of signal propagation, such as Okumura-Hata model, which is one of the most widely used empirical propagation prediction models. It was developed through works of Y. Okumura and M. Hata and is based on the results of extensive measurements in certain urban and sub-urban areas of Japan. Such propagation model is used to predict the signal power of any point on a map, which could be used to assign trust values for SUs. The pattern shown in Figure 4 is typical for a power law based empirical model used in an urban environment. The sector antenna patterns are clearly seen from the shape of the results. The lobes in the vertical pattern of each antenna explain the alternating colours along a radius away from each antenna [31].

The trust value accounts for the density of the surrounding structure of a given SU and the propagation environment in reference to the PBS (source signal) and can be written as:

$$T_{i(t)} = f(d_{i,PBS}(t), h_{i(t)}, h_B, f_0, L, C), \tag{6}$$

where $T_{i(t)} \in \leq T_{i(t)} \leq 1$, $d_{i,PBS}(t)$ denotes the distance between the ith SU and the PBS at time t, $h_i(t)$ denotes the SU height at time t, h_B denotes the PBS height, L denotes the propagation loss, f_0 denotes the central frequency of the signal and C is any physical constant.

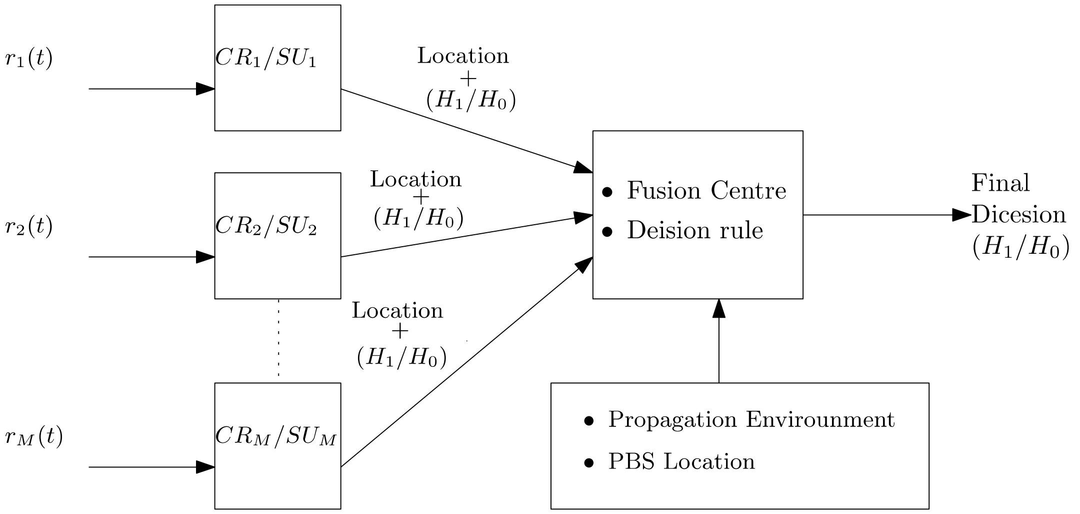


Figure 3. Proposed scheme.

The coverage area of the SBS can be divided into smaller sectors and a trust value is assigned for each sector to represent the environmental propagation in respect to the relevant PBS. The trust value reassesses the sensing data before the fusion process to obtain the global decision. The motivation is to make a comparison between the real sensed signal power, which is received at the SUs, and the expected signal power at each corresponding sector in the coverage area. Trust value contributes to enhance the accuracy of the SBS when the global decision of a particular channel status is calculated.

4.4 Elimination

An SU can be assigned either a low or high trust value. A low trust value indicates that an SU is located in a shadowed area (e.g., highly dense urban area); whereas a high trust value indicates that an SU is located in a less structured region (e.g., LOS propagation is predicted). If an SU is assigned a low trust value, it will be eliminated from subsequent procedures. This step ensures that such an SU does not make any significance when considering a global decision at the SBS. SUs submit the locations and the sensing results simultaneously; therefore, assigning trust values to SUs is time and space dependent. When an SU moves to a new location, a new trust value is assigned which reflects the current location of the SU.

In this paper, we assume that all SUs in the coverage area of the SBS follow the same process. Further steps are taken to SUs, which are assigned a high trust value. SUs measure the received power using energy detection technique, which we briefly discussed in section 3.

SUs submit their local decision to the SBS in a form of hard decision (H_0, H_1). These measurements are further processed at the SBS. Based on the results obtained from the SUs, the SBS determines whether the corresponding channel is free of any primary transmission. We list the detailed procedure in Algorithm 1.

4.5 Proposed Fusion Rule

In cooperative spectrum sensing and in hard combining scheme, SUs send their final one-bit decisions to the SBS. $u_i \in \{0,1\}$ is the binary decision made by the ith SU, which in essence is a logical decision metric. In this context, 0 and 1 indicate the absence and the presence of the primary signal, respectively. There can be a number of fusion rules which are represented by

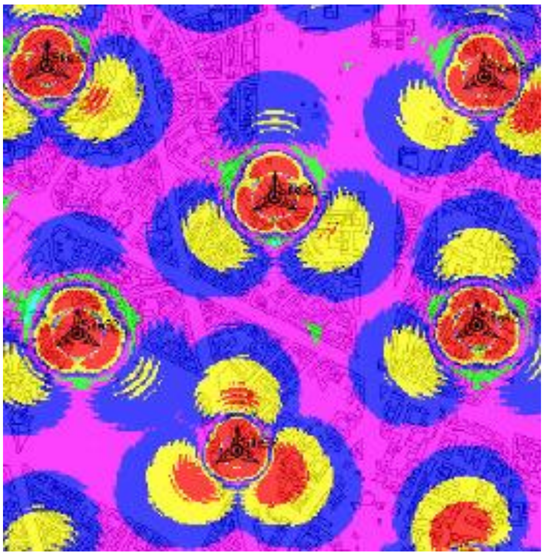


Figure 4. Signal strength from empirical propagation predictions [31].

Algorithm 1 Proposed Spectrum Sensing

```
Initialisation
Number of SUs in the network = M
R ← Empty
for i = 1 : 1 : M do
    Obtain SU's Location
    if ith SU is assigned low trust value then
        Eliminate ith SU from further analysis
    else
        R ← ith SU
        where R is a vector containing all SUs with high
        trust value
    end if
end for
Collect sensing results from SUs in R
Run log likelihood ration test for all SUs in R
Calculate detection and false alarm probability probability
return
```

k-out-of-K rule and for such rule, the overall detection and false alarm probabilities are, respectively:

$$Q_d = \sum_{q=k}^{K} \binom{K}{q} \left\{ \prod_{i=1}^{q} P_{d,i} \times \prod_{j=1}^{K-q} \left(1 - P_{d,i}\right) \right\} \tag{7}$$

and

$$Q_f = \sum_{q=k}^{K} \binom{K}{q} \left\{ \prod_{i=1}^{q} P_{f,i} \times \prod_{j=1}^{K-q} \left(1 - P_{f,i}\right) \right\}; \tag{8}$$

where

$$\binom{K}{q} = \frac{K!/(K-q)!}{q!}. \tag{9}$$

SBS receives decisions from M SUs, decide H_1 if any of the total M individual decisions is H_1 and decides H_0 otherwise. This fusion rule is known as the OR-rule or 1-out-of-M rule, while AND rule corresponds to the case where $k = K$. SBS receives decisions from M SUs and decides H_1 if all of the total M individual decisions are H_1 and decides H_0 otherwise. The global probabilities of false alarm and detection for the two fusion rules can be obtained as:

OR fusion rule:

$$Q_{d,or} = 1 - (1 - P_d)^M \tag{10}$$

and

$$Q_{f,or} = 1 - (1 - P_f)^M \tag{11}$$

AND fusion rule:

$$Q_{d,and} = (P_d)^M \tag{12}$$

and

$$Q_{f,and} = (P_f)^M. \tag{13}$$

Fusion of incoming local decisions and decisions that are made at the SBS are considered in this paper. In the scenario discussed here, SUs could make only hard decisions, such that u_i could take only two values 0 or 1 based on its local observation $u_i \in \{0,1\}$. All the local detector SUs observe the same channel at the same time. Each SU makes a local binary decision u_i, where $\{i = 1,..., M\}$ based on the local observation. The SBS produces the global decision $u_o \in \{0,1\}$. This problem is known as the binary hypothesis test, since the system chooses, between two hypotheses, where H_0 and H_1 are the noise only hypothesis and the signal plus noise hypothesis, respectively. The optimum fusion rule for this problem is given by the likelihood ratio test (LRT) as:

$$\frac{\Pr(u_1, u_2, ..., u_M | H_1)}{\Pr(u_1, u_2, ..., u_M | H_0)} \underset{u_o=0}{\overset{u_o=1}{\gtrless}} \eta, \tag{14}$$

where η is a threshold which is determined by the specified values of $P_{d,i}$ and $P_{f,i}$. Next, we assume that $P_{d,i} \geq P_{f,i}$, where $\{i = 1,..., M\}$. This assumption is common in CCR network sensing scenarios. We also make the following definitions:

$$P_{f,i} = \Pr(u_i = 1 | H_0) \tag{15}$$

and

$$P_{d,i} = \Pr(u_i = 1 | H_1), \tag{16}$$

where u_o is the global decision at the SBS. Given this assumption, the optimum fusion rule can be written as:

$$\frac{\Pr(u_1,u_2,...,u_M|H_1)}{\Pr(u_1,u_2,...,u_M|H_0)}=\prod_{i=1}^{M}\frac{\Pr(u_i=j|H_1)}{\Pr(u_i=j|H_0)}=\prod_{S_1}\frac{\Pr(u_i=1|H_1)}{\Pr(u_i=1|H_0)}\times\prod_{S_0}\frac{\Pr(u_i=0|H_1)}{\Pr(u_i=0|H_0)}$$

$$=\prod_{S_1}\frac{P_{d,i}}{P_{f,i}}\times\prod_{S_0}\frac{1-P_{d,i}}{1-P_{f,i}}, \quad (17)$$

where S_j is the set of all decisions made at the SBS that are equal to j, $j=0,1$. The fusion rule that minimizes the probability of false alarm and maximizes the probability of detection is given by:

$$\prod_{S_1}\frac{P_{d,i}}{P_{f,i}}\times\prod_{S_0}\frac{1-P_{d,i}}{1-P_{f,i}}\underset{u_o=0}{\overset{u_o=1}{\gtrless}}\eta. \quad (18)$$

So far, we have discussed the fusion rules for the binary hypothesis problem. Next, we include the case in which the SU is assigned a trust value which represents the signal strength in its respective region. Trust values are modeled as the probability of an SU to be located in a region of acceptable reception, e.g., $T_i=j$, and $j\in 0\le j\le 1$, where T_i is spatially independent and $j=0$ represents the respective SU location being in a high shadowed area, while $j=1$ indicates that a user is located within a line of sight in respect to the sensed signal (source signal). These trust values are transmitted to the fusion center for further processing to reach the global decision. The fusion rule in this case is given by the LRT in (19), since it indicates for each value of Ti the likelihood of H_1 versus the likelihood of H_0 and can be expressed as:

$$\frac{\Pr(T_1,T_2,...,T_M|H_1)}{\Pr(T_1,T_2,...,T_M|H_0)}\times\frac{\Pr(u_1,u_2,...,u_M|H_1)}{\Pr(u_1,u_2,...,u_M|H_0)}\underset{u_o=0}{\overset{u_o=1}{\gtrless}}\eta. \quad (19)$$

Now, let us define the following probabilities:

$$\alpha_{i,j}=Pr\{T_i=j\,|H_0\} \quad (20)$$

$$\beta_{i,j}=Pr\{T_i=j\,|H_1\}. \quad (21)$$

The ratio $\frac{\Pr(T_1,T_2,...,T_M|H_1)}{\Pr(T_1,T_2,...,T_M|H_0)}$ in Equation (19) can be expressed as:

$$\frac{\Pr(T_1,T_2,...,T_M|H_1)}{\Pr(T_1,T_2,...,T_M|H_0)}=\prod_{i=1}^{M}\frac{\Pr(T_i|H_1)}{\Pr(T_i|H_0)}=\prod_{i=1}^{M}\prod_{S_j}\frac{\Pr(T_i=j|H_1)}{\Pr(T_i=j|H_0)}. \quad (22)$$

Substituting (20) and (21) in (22), we obtain the following expression:

$$\frac{\Pr(T_1,T_2,...,T_M|H_1)}{\Pr(T_1,T_2,...,T_M|H_0)}=\prod_{j=o}^{j-1}\prod_{S_j}\left(\frac{\beta_{ji}}{\alpha_{ji}}\right). \quad (23)$$

Subsequently, by substituting (23) and (18) in (19), we obtain the following fusion rule:

$$\prod_{j=o}^{j-1}\prod_{S_j}\left(\frac{\beta_{ji}}{\alpha_{ji}}\right)\times\prod_{S_1}\frac{P_{d,i}}{P_{f,i}}\times\prod_{S_0}\frac{1-P_{d,i}}{1-P_{f,i}}\overset{u_o=1}{\underset{u_o=0}{\gtrless}}\eta, \tag{24}$$

and by taking the logarithm of both sides, we obtain the optimum fusion rule that minimizes the false alarm and maximize the probability of detection as:

$$\sum_{j=o}^{j-1}\sum_{S_j}\log\left(\frac{\beta_{ji}}{\alpha_{ji}}\right)+\sum_{S_1}\log\frac{P_{d,i}}{P_{f,i}}+\sum_{S_0}\log\frac{1-P_{d,i}}{1-P_{f,i}}\overset{u_o=1}{\underset{u_o=0}{\gtrless}}\log\eta. \tag{25}$$

5. Analysis and Results

In order to evaluate the performance of our scheme, analytical results are given in this section. In our analysis, it is assumed that the SBS is aware of the relevant primary network parameters, as well as the locations of SUs and PBS and the trust value can be calculated. In the analysis, we set the number of cooperative SUs to be 30. For the analytical results, it is reasonable that we compare our proposed scheme with the conventional cooperative spectrum sensing schemes. The comparison is presented in Figure 5. We varied the signal to noise ratio (SNR) form (-2dB) to (2dB). We consider two false alarm probability P_f values which are set to be 0.1 and 0.2, while the SUs which are located in high trust value are set at 0.75. Figure 5 shows the improved performance of our proposed scheme when eliminating the SUs which are considered to be located in high shadowed areas, with a percentage of 25% of all participating SUs. Because these SUs are eliminated from further processing, they have no impact on the final global decisions. It is clear for both values of false alarm probabilities that the probability of detection P_d increased when we apply our proposed scheme. Results also indicate a slight improvement in terms of required average SNR for detection.

The results in Figure 6 show the ROC performance comparison of the proposed location-aware and conventional (or the case where location and propagation models are not considered) cooperative spectrum sensing schemes when T = 0.78 and T = 0.6. T = 0.78 indicate that 22% of the SU are located in a highly shadowed areas. These SUs are eliminated from further processing at the SBS. The location-aware scheme slightly outperforms the conventional scheme when most of the SUs are located in the same environment. However, Figure 6 shows that the performance has improved further when T = 0.6, which indicates that 40% of the SUs are located in unreliable locations.

Table 2. Sensing procedure comparison.

	Proposed Scheme	Conventional Schemes
SUs Make Local Decisions	✓	✓
SUs Send Decision to SBS	✓	✓
SUs Send Geo-Location to SBS	✓	
SBS Calculate Distance from PBS to SUs	✓	
SBS Calculate Channel Condition	✓	
SBS Calculates Trust value for each SU	✓	
SBS Calculates Global Decision	✓	✓

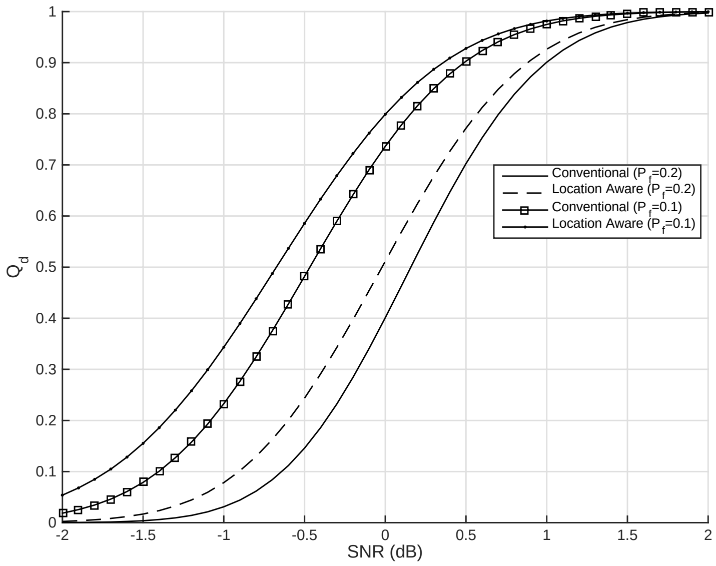


Figure 5. Probability of detection comparison of proposed location-aware scheme and conventional hard combining scheme for different SNR when false alarm probability constraint is 0.1 and 0.2.

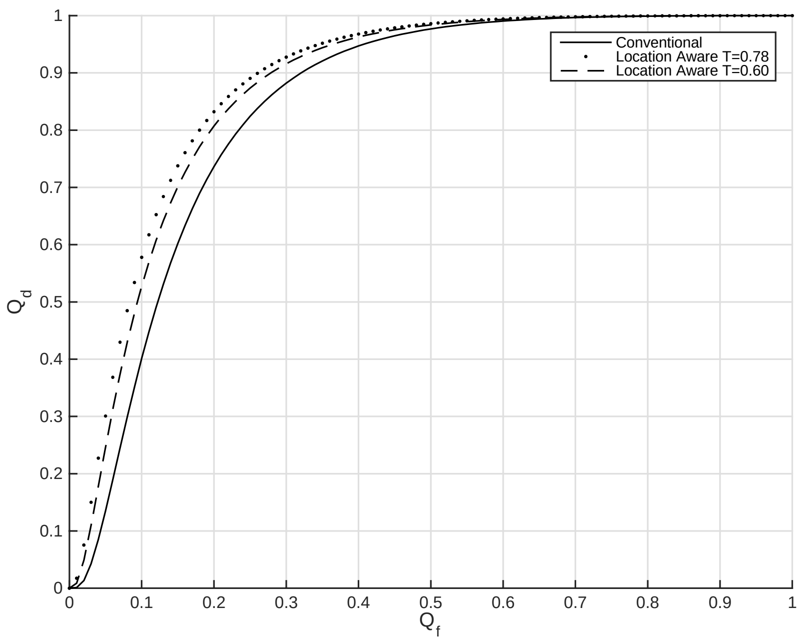


Figure 6. ROC comparison of proposed location-aware scheme and conventional hard combining scheme under Gaussian channel when the number of cooperative users = 30 for different trust values.

Figure 7 depicts the ROC analytical curves using the proposed scheme when the number of cooperative users is 30 and 35% of the SUs are located in highly shadowed areas. The high value of sensing results means that most of the participating CRs are located nearby the source signal. On the other hand, since 35% of the CRs are located too far from the source signal and/or NLOS is predicted, because they are located in highly shadowed areas, those CRs can not be considered as valuable sources of information. Therefore, they are eliminated from further action. It is evident from Figure 7 that the proposed cooperative sensing scheme provides better performance than the conventional hard combination scheme. When T=0.65, for the detection probability of 0.9, the false alarm probability of the proposed scheme is 0.25, while for the conventional scheme it is 0.32.

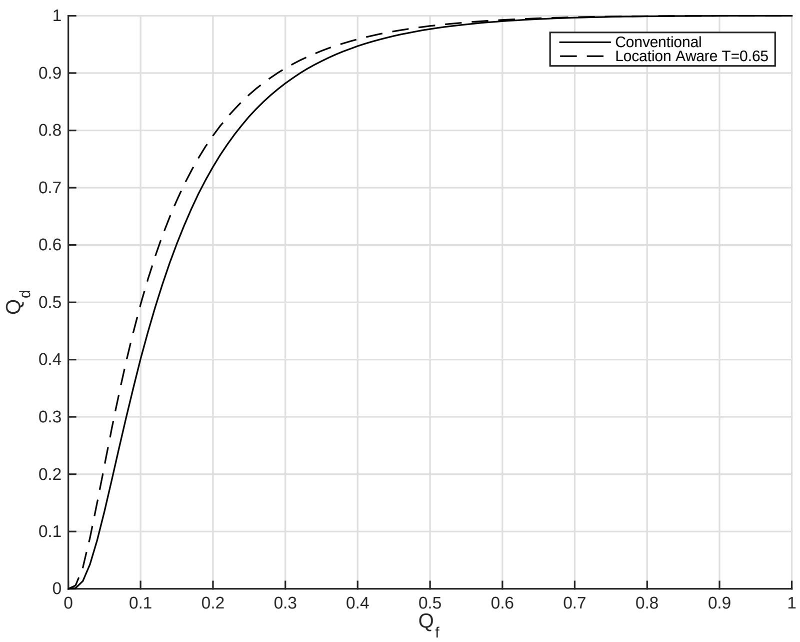


Figure 7. ROC comparison of proposed location-aware scheme and conventional hard combining scheme under Gaussian channel when the number of cooperative users = 30 and 35% of the SUs are located in highly shadowed areas.

In Figure 8, we plot the probability of detection against the SNR. The figure presents the probabilities of detection for different numbers of cooperative cognitive radios in the network.

It is evident that the detection improves with increased number of CRs, since more accurate results mean better performance for the network. The number of CRs is typically large in the case of urban networks. However, the proposed scheme can eliminate the CRs with low trust value from participating in the cooperative sensing. The proposed scheme does not only improve performance of detection, but also reduces sensing time.

Cooperative spectrum sensing may become impractical in CRNs with a large number of SUs, because in a time slot only one SU sends its local decision to the SBS in order for the decisions to be separated easily. Hence, it may make the whole sensing time intolerably long. The scheme proposed here does not take into account the users that are located in low trust value regions, therefore it minimizes the number of participating SUs in a selective manner. Consequently, the processing time for the global decision at the SBS will be minimized while not compromising spatial diversity. This implies that SBSs have the incentive to adapt the proposed sensing decision method, since it can lead to achieve higher reliability and lower sensing time. The fundamental differences between our proposed scheme and the conventional methods are shown in Table 2.

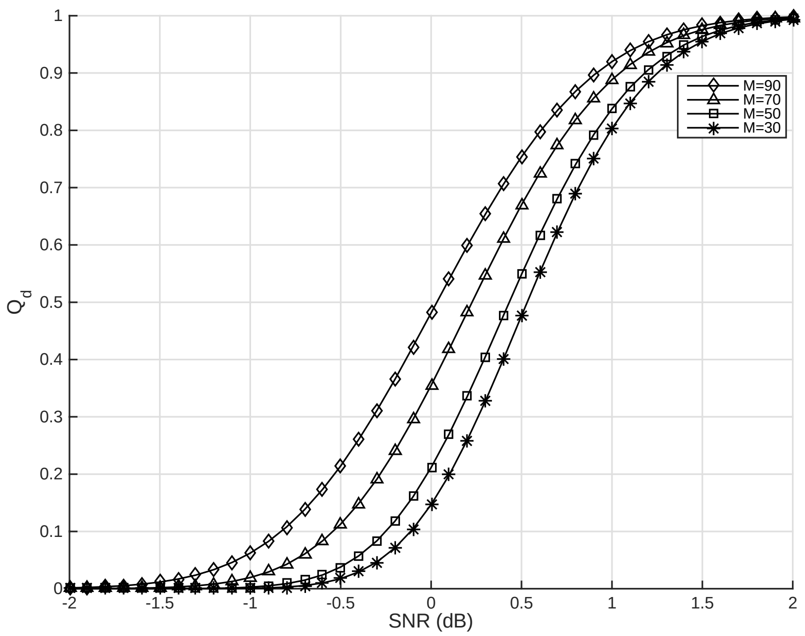


Figure 8. ROC of proposed location-aware scheme with different numbers of cognitive radio users under Gaussian channel and 35% of the SUs are located in highly shadowed areas.

6. CONCLUSIONS

We have studied the performance of cooperative cognitive spectrum sensing with energy detection in CR networks. Location-aware cooperative spectrum sensing has been investigated. We have derived the optimum fusion rule, as well as the probability of detection, taking location reliability into consideration. The proposed scheme has proved to exhibit better ROC, especially in highly shadowed regions (e.g., under NLOS propagation conditions). Analytical results of the proposed location-aware scheme show an improved performance over the conventional hard combining schemes (e.g., [32]), highlighting the requirements of location knowledge in CRNs, especially in urban environments. Since this sensing accuracy is mainly related to the signal propagation environments, the more accurate the propagation models are, the better the expected performance will be from our proposed scheme. Moreover, for a cognitive radio network, high probability of detection results in less interference to the primary network, which means more capacity and more offered service at high quality. A major issue concerning the practical implementation of the proposed scheme is the availability of complete statistical information corresponding to source signal parameters, particularly their variation with distance. However, lack of spectrum resources encourages the adoption of new ways of sharing, including sharing of specific data related to the incumbent operators.

There are several natural directions suggested by our paper. The most obvious one is to utilize the eliminated CRs from the first step of the cooperative sensing. For example, it would be interesting to develop some more complex schemes of spectrum sensing, e.g., assign the eliminated CRs to sense different channels which are in LOS and/or in close proximity to different source signals. This could improve the efficiency of sensing, not only by sensing more channels simultaneously but also with high accuracy. Moreover, to gain further understating of our proposed scheme, the sensing time performance could be evaluated.

In the case of Universal Mobile Telecommunications System (UMTS), the transmitting power is adapted to the propagation conditions. The transmitting power is always selected to be only as high as necessary for adequate connection quality. Moreover, each service supported by UMTS networks requires specific threshold values and the network behavior and size change with traffic. Data transmission adds yet another dimension of complexity. This makes detecting UMTS signals much more difficult than in the case of other technology; e.g., Global System for Mobile Communications (GMS). Therefore, it would be very useful to conduct a study that specifically addresses UMTS networks.

REFERENCES

[1] A. Palaios, J. Riihijarvi, P. Mahonen, V. Atanasovski, L. Gavrilovska, P. Van Wesemael, A. Dejonghe and P. Scheele, "Two Days of Spectrum Use in Europe," in 7th International ICST Conference on Cognitive Radio Oriented Wireless Networks and Communications (CROWNCOM), pp. 24–29, 2012.

[2] M. Hoyhtya, M. Matinmikko, X. Chen, J. Hallio, J. Auranen, R. Ekman, J. Roning, J. Engelberg, J. Kalliovaara, T. Taher, A. Riaz and D. Roberson, "Measurements and Analysis of Spectrum Occupancy in the 2.3 x2013; 2.4 GHz Band in Finland and Chicago," in 9th International Conference on Cognitive Radio Oriented Wireless Networks and Communications (CROWNCOM), pp. 95–101, 2014.

[3] B. Wang and K. Liu, "Advances in Cognitive Radio Networks: A Survey," IEEE Journal of Selected Topics in Signal Processing, vol. 5, pp. 5–23, February 2011.

[4] P. Kolodzy, "Spectrum Policy Task Force Report," Federal Communications Commission, www.fcc.gov/sptf, 2002.

[5] A. Wyglinski, M. Nekovee and Y. Hou, Cognitive Radio Communications and Networks: Principles and Practice, Academic Press, 2010.

[6] N. Pappas and M. Kountouris, "Throughput of a Cognitive Radio Network under Congestion Constraints: A Network-level Study," pp. 162–166, June 2014.

[7] I. Akyildiz, W.-Y .Lee, M. C. Vuran and S. Mohanty, "Asurvey on Spectrum Management in Cognitive Radio Networks," IEEE Communications Magazine, vol. 46, pp. 40–48, June 2008.

[8] S. Haykin, "Cognitive Radio: Brain-Empowered Wireless Communications," IEEE Journal on Selected Areas in Communications, vol. 23, pp. 201–220, Feb. 2005.

[9] S. Atapattu, C. Tellambura and H. Jiang, "Spectrum Sensing via Energy Detector in Low SNR," in IEEE International Conference on Communications (ICC), pp. 1–5, 2011.

[10] I. F. Akyildiz, W.-Y. Lee, M. C. Vuran and S. Mohanty, "Next Generation/Dynamic Spectrum Access/Cognitive Radio Wireless Networks: A Survey," Computer Networks, vol. 50, no. 13, pp. 2127–2159, 2006.

[11] L. Wei and O. Tirkkonen, "Statistical Test for Multiple Primary User Spectrum Sensing," pp. 41–45, 2011.

[12] E. Axell, G. Leus, E. G. Larsson and H. Poor, "Spectrum Sensing for Cognitive Radio: State-of-the-art and Recent Advances," in IEEE Signal Processing Magazine, vol. 29, no. 3, pp. 101–116, 2012.

[13] A. Ghasemi and E. Sousa, "Spectrum Sensing in Cognitive Radio Networks: Requirements, Challenges and Design Trade-offs," in IEEE Commu nications Magazine, vol. 46, no. 4, pp. 32–39, 2008.

[14] N. Kundargi and A. Tewfik, "Doubly Sequential Energy Detection for Distributed Dynamic Spectrum Access," in IEEE International Conference on Communications (ICC), pp. 1–5, May

2010.

[15] S. Atapattu, C. Tellambura and H. Jiang, "Energy Detection of Primary Signals over Fading Channels," in International Conference on Industrial and Information Systems (ICIIS), pp. 118–122, Dec. 2009.

[16] S. Atapattu, C. Tellambura and H. Jiang, "Performance of an Energy Detector over Channels with Both Multipath Fading and Shadowing," in IEEE Transactions on Wireless Communications, vol. 9, pp. 3662–3670, Dec. 2010.

[17] X. Liu, C. Zhang and X. Tan, "Double-threshold Cooperative Detection for Cognitive Radio Based on Weighing," in 6th International ICST Conference on Communications and Networking in China (CHINACOM), pp. 205–209, 2011.

[18] K. Ben Letaief and W. Zhang, "Cooperative Communications for Cognitive Radio Networks," Proceedings of the IEEE, vol. 97, no. 5, pp. 878–893, May 2009.

[19] Q. Zou, S. Zheng and A. Sayed, "Cooperative Spectrum Sensing via Sequential Detection for Cognitive Radio Networks," in IEEE 10th Workshop on Signal Processing Advances in Wireless Communications (SPAWC), pp. 121–125, 2009.

[20] H. Mu and J. Tugnait, "Joint Soft-decision Cooperative Spectrum Sensing and Power Control in Multi-band Cognitive Rradios," Processing of the IEEE Transactions on Signal, vol. 60, no. 10, pp. 5334–5346, 2012.

[21] Y. Tani and T. Saba, "Quantization Scheme for Energy Detector of Soft Decision Cooperative Spectrum Sensing in Cognitive Radio," in IEEE GLOBECOM Workshops (GC Wkshps), pp. 69–73, 2010.

[22] S. Althunibat, M. Di Renzo and F. Granelli, "Towards Energy-efficient Cooperative Spectrum Sensing for Cognitive Radio Networks: An Overview," Telecommunication Systems, vol. 59, no. 1, pp. 77–91, 2015.

[23] S. Althunibat and F. Granelli, "An Objection-based Collaborative Spectrum Sensing for Cognitive Radio Nnetworks," IEEE Communications Letters, vol. 18, no. 8, pp. 1291–1294, 2014.

[24] B. Shen, L. Huang, C. Zhao, K. Kwak and Z. Zhou, "Weighted Cooperative Spectrum Sensing in Cognitive Radio Networks," in 3rd International Conference on Convergence and Hybrid Information Technology (ICCIT), vol. 1, pp. 1074–1079, 2008.

[25] S. Chaudhari, J. Lunden, V. Koivunen and H. Poor, "Cooperative Sensing with Imperfect Reporting Channels: Hard Decisions or Soft Decisions?," Processing of the IEEE Transactions on Signal, vol. 60, no. 1, pp. 18–28, 2012.

[26] S. A. Alvi, M. S. Younis, M. Imran, M. Guizani et al., "A Near-Optimal LLR-based Cooperative Spectrum Sensing Scheme for Crahns," in IEEE Transactions on Wireless Communications, vol. 14, no. 7, pp. 3877– 3887, 2015.

[27] D. Hamza, S. Aissa and G. Aniba, "Equal Gain Combining for Cooperative Spectrum Sensing in Cognitive Radio Networks," in IEEE Transactions on Wireless Communications, vol. 13, no. 8, pp. 4334–4345, 2014.

[28] X. Li, "RSS-based Location Estimation with Unknown Pathloss Model," in IEEE Transactions on Wireless Communications, vol. 5, no. 12, pp. 3626–3633, 2006.

[29] G. E. Athanasiadou, A. R. Nix and J. P. McGeehan, "A Microcellular Ray-tracing Propagation Model and Evaluation of Its Narrow-band and Wide-band Predictions," IEEE Journal on Selected Areas in Communications, vol. 18, no. 3, pp. 322–335, 2000.

[30] Y. Qi, H. Kobayashi and H. Suda, "Analysis of Wireless Geolocation in a Non-line-of-sight Environment," in IEEE Transactions on Wireless Communications, vol. 5, no. 3, pp. 672–681, 2006.

[31] M. Coinchon, A.-P. Salovaara and J.-F. Wagen, "The Impact of Radio Propagation Predictions on Urban UMTS Planning," Proceedings of the IEEE 2002 International Zurich Seminar on

Broadband Communications, Access, Transmission, Networking, pp. 32-1 - 32-6, Feb. 2002.

[32] S. Atapattu, C. Tellambura and H. Jiang, "Energy Detection Based Cooperative Spectrum Sensing in Cognitive Radio Networks," in IEEE Transactions on Wireless Communications, vol. 10, pp. 1232–1241, Apr. 2011.

ملخص البحث:

يُعدّ استشعار الطيف من تقنيات التمكين الأساسية لشبكات الراديو الإدراكية (*CRNs*). والهدف الرئيسي من استشعار الطّيف هو توفير المزيد من فرص الوصول إلى الطيف لمستخدمي الراديو الإدراكيّ دون التداخل مع عمليات الشبكة المرخّصة. ويمكن للقرارات المتعلقة باستشعار الطيف أن تقود إلى استشعار غير صحيح ذي أداء متدنٍّ بسبب الاضمحلال أو التظليل أو التداخلات الأخرى الناجمة إمّا عن عدم تجانس التضاريس أو البنية الحضرية الكثيفة.

ومن أجل تحسين قرارات استشعار الطيف، نقترح في هذه الورقة نظاماً تعاونياً لاستشعار الطيف. ويأخذ النظام المقترح بعين الاعتبار ظروف الانتشار، مثل: تفاوت التضاريس وشدّتها والبنية الحضرية بين نقطتين بالنسبة لانتشار الموجات. كذلك، تمّ اشتقاق قاعدة الاندماج الأمثل التي تفسّر موثوقية الموقع للمستخدمين الثانويّين. وتبين النتائج التحليلية أن النظام المقترح يتفوق بعض الشيء على الطّرق التقليدية المستخدمة في الاستشعار التعاوني للطّيف.

Jordanian Journal of Computers and Information Technology (JJCIT), Vol. 2, No. 2, August 2016.

A HARDWARE-EFFICIENT BLOCK MATCHING UNIT FOR H.265/HEVC MOTION ESTIMATION ENGINE USING BIT-SHRINKING

Osama M. F. Abu-Sharkh[1] and Esam A. AlQaralleh[2]

Department of Computer Engineering, Princess Sumaya University for Technology, Amman, Jordan.
{osama[1], qaralleh[2]}@psut.edu.jo

(Received: 28-Dec.-2015, Revised: 22-Mar.-2016, Accepted: 10-Apr.-2016)

ABSTRACT

The main objective of this work is to enhance the processing performance of the recently introduced video codec H. 265/HEVC. Since most of the computations of H. 265/HEVC still occur in the motion estimation engine which is inherited from its predecessor H.264/AVC, we propose a bit-shrinking approach with a modified logic functionality to design an efficient and simplified block matching unit that replaces the already used Sum of Absolute Differences (SAD) unit. The hardware complexity of the proposed unit itself is reduced and the number of its generated output bits is reduced as well which in turn simplifies all the subsequent units of motion estimation. The hardware complexity, the consumed power and the processing delay of the motion estimation engine are therefore reduced significantly with only marginal deterioration in both the bit-rate and the peak-signal-to-noise-ratios (PSNR) of the tested High Definition (HD) and Ultra-High Definition (UHD) H.265/HEVC compressed videos. We simulate our design using HM16.6 and perform system logic synthesis using Synopsys's Design Compiler, targeting ASIC, for evaluation purposes.

KEYWORDS

H.265 / HEVC, Ultra-High Definition, Motion Estimation, Video Coding, Bit-Truncation.

1. INTRODUCTION

For almost a decade, H.264/AVC [1] has been the *de facto* standard for video coding. Its impressive performance opened wide doors for watching and exchanging videos almost everywhere. In spite of the fact that H.264 can handle High Definition (HD) videos, the sizes of these videos using H.264/AVC are still a concern, especially when using smart handheld devices with limited storage and power constraints. Recently, H.265/HEVC [2] has been introduced in the literature to provide better video coding performance than the legacy H.264/AVC. The former can provide up to twice the compression ratio of the latter, while maintaining the same video quality. Rearticulating, the former can provide much better video quality than the latter for the same video compression ratio. This huge increase in performance is due to the many enhanced techniques and methodologies that have been introduced in H.265/HEVC. Some of these enhancements have tackled the block matching criterion of motion estimation, the powerful engine of video coding. Although, block matching still relies on the use of the Sum of Absolute Differences (SAD) in H.265/HEVC which is inherited from H.264/AVC, the maximum block size has been enlarged in the former to 64x64 pixels instead of 16x16 pixels. Thus, the number of computations and consequently the number of the used SAD components and its constructing logic gates vastly increased in H.265/HEVC. Hence, the hardware complexity, the processing delay and the power consumption of the logic gates increased vastly as well. Taking the processing delay as an example, Grois et al. conducted experiments in a recent work [3] to compare the performance of H.264/AVC, VP9 [4] and H.265/HEVC. They reported that the typical total encoding time of VP9 is around 130 times higher than the total encoding time of H.264/AVC and 7.35 times lower than the total encoding time of H.265/HEVC for the same PSNR value. This means that the total encoding time of H.265/HEVC is 955.5 times higher than that of H.264/AVC for the same PSNR value, which is actually a huge increase in the encoding time. The motion estimation engine by itself consumes around 60% to 80% of the total encoding time of H.264/AVC [5], while it consumes around 80% of the total encoding time of H.265/HEVC [6]. Thus, a pragmatic strategy to elevate the performance of H.265/HEVC relies on enhancing the performance of the core component of the motion estimation engine which is the block matching unit.

The main focus of this work is the block matching unit of the motion estimation engine. We propose in this paper a simplified and hardware-efficient block matching unit that replaces SAD. The framework of the proposed work is the motion estimation engine of the state-of-the-art video codec, H.265/HEVC. The main contributions and the differences from others in the literature are explained thoroughly in the following section.

The rest of the paper is organized as follows. We survey some of the related work in the literature and clarify our contribution in section 2. We then discuss the proposed designs of the block matching unit in section 3. We provide numerical analysis and evaluations with discussion in section 4 and finally conclude the paper in section 5.

2. LITERATURE REVIEW AND CONTRIBUTION

Many works in the literature have considered simplifying either the software and/or the hardware design of motion estimation in order to enhance the performance of video processing. Fast matching algorithms such as three steps search [7], four steps search [8], hexagon-based search [9], diamond search [10] and adaptive rood pattern search [11] tend to shrink the number of matched macro-blocks immensely. Most of these algorithms are only suitable for software implementations which are not as efficient as hardware implementations for real-time applications. Online arithmetic [12] and saturation arithmetic [13] are used to reduce the computational complexity of SAD. Nevertheless, the tree-adder implementation, where the SAD computations are performed, is still either time consuming or hardware costly. Vanne et al. performed in [14] arithmetic operations accompanied with several early termination mechanisms and sophisticated SAD computation control. Another approach based on SAD [15] takes the difference pixel count (DPC) as the selection criterion. Yeo et al. [16] use an XOR function instead of adders to simplify the matching criterion. One of the most recent promising approaches is to reduce the pixel resolution from eight bits to fewer bits. The works introduced in [17]-[20] use a one-bit transform by converting video frames into a single-bit-plane. On the other hand, the works introduced in [21]-[22] follow a bit-truncation approach where least significant bits are eliminated to simplify the hardware. One-bit transform and bit-truncation approaches deteriorate both the compression ratio and the video quality. The amount of deterioration in the latter depends on the number of truncated bits. Recently, Manjunatha and Sainarayanan recommended in [23] the use of a 1-bit full adder which consists of XOR, AND and OR gates instead of the commonly used 1-bit full adder which consists of XOR and NAND gates for performing SAD. In their work, they showed enhancements in the consumed power, latency and area when using the former rather than the latter. Following a different approach, we introduced in [24] a modified XOR function that replaces conventional SAD of the motion estimation engine of H.264/AVC. We showed enhancements over many proposed designs in the literature. All of the aforementioned authors who evaluated their works based on video coding, adopted the motion estimation engine of H.264/AVC with a maximum block size of 16x16 and low quality videos, CIF and QCIF in their evaluations.

Jordanian Journal of Computers and Information Technology (JJCIT), Vol. 2, No. 2, August 2016.

Some recent works, which consider the enhancements of the motion estimation engine of H.265/HEVC, have also been proposed in the literature. We discuss some of the efforts which consider hardware implementations as follows. Sanchez et al. evaluated in [25] the use of the Multi-Point Diamond Search (MPDS) algorithm [26] in H.265/HEVC. They found that it is more hardware-friendly than the Enhanced Predictive Zonal Search (EPZS), which is implemented in the standard, on the expense of small amounts of deterioration in the video quality and compression ratio. Sinangil and Sze proposed in [27] a new hardware-aware search algorithm for HEVC motion estimation. They reported enhancements in area and bandwidth when using their algorithm. Jaja et al. proposed in [28] two fast motion estimation algorithms based on the structure of the triangle and the pentagon for H.265/HEVC. In their experimental evaluations, they found that the proposed algorithms can offer up to 63% and 61.9% speed-up in run-time when compared with the original algorithms of the standard. Medhat et al. proposed in [29] a Fast Center Search Algorithm (FCSA) for H.265/HEVC. They indicated that FCSA reaches average time saving ratio up to 40% for HD video sequences with insignificant loss in PSNR and compression ratio. Miyazawa et al. introduced in [30] a complete hardware implementation for H.265/HEVC and evaluated its efficiency in processing videos in real-time applications. They compared their implementation with a professional-use H.264/AVC video encoder available in the market. They were able to encode HD videos at 60fps in real-time. Pastuszak and Trochimiuk proposed in [31] a high-throughput motion estimation system to process Ultra-High Definition (UHD) videos in H.265/HEVC. The system embeds two parallel processing paths for the integer-pel and the fractional-pel motion estimation. Their synthesis showed that the system is able to encode UHD videos at 30fps with only small deteriorations in PSNR and compression ratio. Ye et al. proposed in [32] a parallel clustering tree search (PCTS) algorithm for integer-pel motion estimation that processes the prediction units (PU) simultaneously with a parallel scheme. The hardware implementation of PCTS can support quad-full HD (QFHD) videos at 30fps in real-time. All of the aforementioned efforts still rely on the use of the block matching unit, conventional SAD in designing the algorithms to enhance the motion estimation of H.265/HEVC.

In this paper, we introduce several enhancements on our previous work [24]. They are mentioned in the following. Our new modified block matching unit is implemented in the motion estimation engine of H.265/HEVC which has a different hardware-efficiency with the increased block size than the one used by H.264/AVC. We also introduce in this paper a new bit-shrinking approach to reduce the number of generated output bits of the matching unit which is reflected on all subsequent stages of the motion estimation engine. We perform system logic synthesis using Synopsys's Design Compiler [33], targeting ASIC, to evaluate our design and compare it with the conventional SAD and other works in the literature. We consider the number of gates, the consumed power and the processing delay as performance metrics for evaluation purposes. The obtained results show the superiority of our design in all performance metrics. We also run extensive simulations using HM16.6 [34] to measure the video quality and the compression ratio. We apply our simulations on both HD and UHD videos for evaluation purposes and consider a block size of 64x64.

It is worthy to mention that the proposed block matching unit can be utilized by many motion estimation algorithms in the literature, such as the ones proposed in [25]-[32], to replace conventional SAD and hence enhance performance.

3. PROPOSED BLOCK MATCHING UNIT

Let us first give a brief description of the sum of absolute differences operation. SAD is a measure of similarity between a block of pixels of the present frame and a block of pixels of a previous reference frame of a video. It estimates motion between the considered frames to remove redundant information and consequently reduces the sizes of the videos. Considering NxN as the macro-block size, $p_k(i,j)$ as the current pixel of a macro-block, $p_{k-1}(i+u,j+v)$ as the candidate pixel of a macro-block of a reference frame and $[-q, q-1]$ as the search range, SAD is defined by the following equation.

$$SAD(u,v) = \sum_{i=0}^{N-1}\sum_{j=0}^{N-1} |p_k(i,j) - p_{k-1}(i+u, j+v)|; \qquad (1)$$

where $-q \le u, v \le q-1$.

The absolute difference function in Equation (1) can be described as:

$$|X - Y| = \begin{cases} X + \bar{Y} + 1 & if\ MSB = 0 \\ \bar{X} + Y + 1 & if\ MSB = 1 \\ 0 & if\ X = Y \end{cases} \qquad (2)$$

The sum of absolute difference function involves the comparison of two pixels. These pixels have unsigned representations of luminous values. The unsigned nature of these pixels shall burden the system. To avoid this, the sum of absolute difference function is designed in several ways in hardware. One way is by subtracting two unsigned numbers and then making a decision about the obtained result by converting the negative sum into positive as shown in Figure 1 with the aid of an XOR gate. Another design is achieved by subtracting the first number from the second one and also the second number from the first one, simultaneously, and then selecting the positive result via a multiplexer. The former implementation encounters longer delay than the latter, since the critical path goes through two adders in the former while it goes only through one adder and one multiplexer in the latter. Note that the two adders in the second design operate in parallel.

There are several hardware implementations for the full adder. The basic, simplest and mostly used implementation is the Ripple Carry Adder (RCA). As the name indicates, the carry ripples from the n^{th} bit to the next one and so on until it reaches the most significant bit. Thus, the most significant bit of the output cannot be calculated until all the preceding bits are calculated one after another. It can be deduced that the more the number of calculated bits is, the more is the delay in generating the final result. Another implementation of a full adder is the Carry Look Ahead Adder (CLA). The latter is faster than the RCA, but more complex, which leads to a larger implementation area and more power consumption.

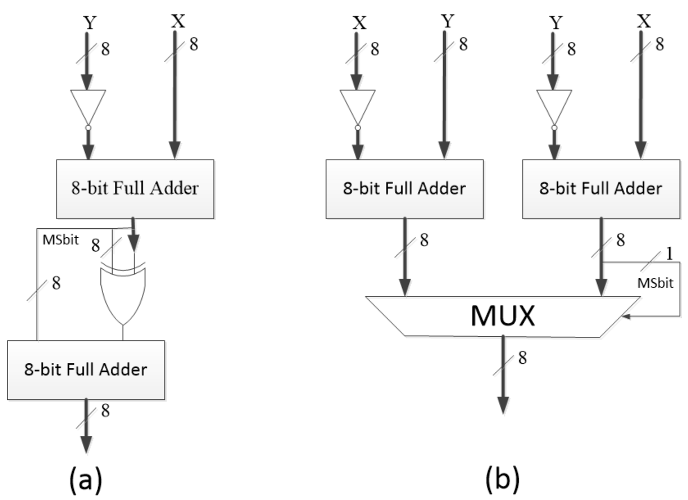


Figure 1. Two different designs for the sum of absolute difference circuit.

Jordanian Journal of Computers and Information Technology (JJCIT), Vol. 2, No. 2, August 2016.

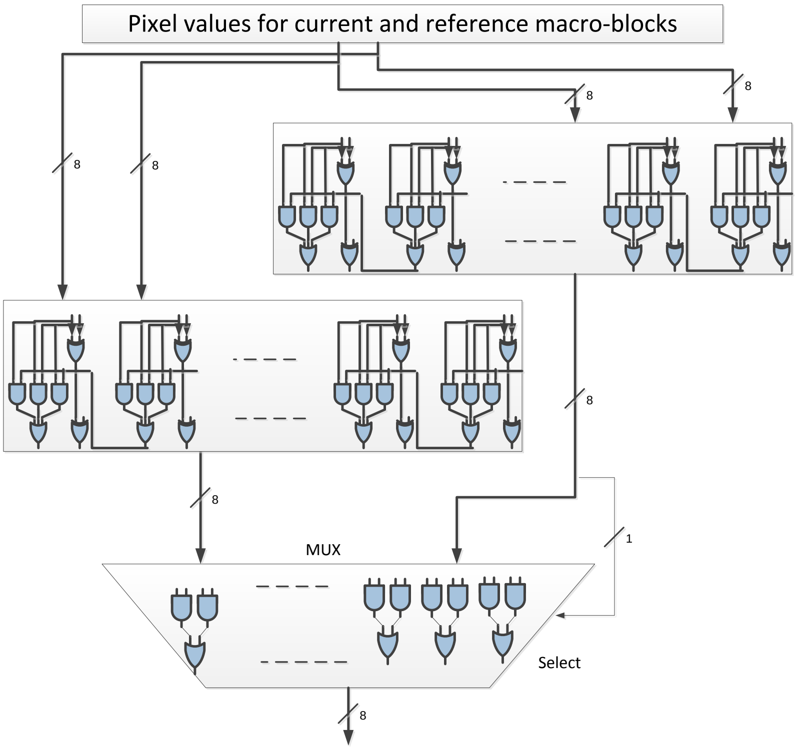


Figure 2. Sum of absolute difference circuit using 8-bit full adders and a multiplexer.

Considering the circuit shown in Figure 1. (b), the sum of absolute difference function for one pixel matching is implemented using two 8-bit full adders and one multiplexer. These units are built from many logic gates as shown in Figure 2. Each macro-block matching consists of many SAD operations. Since very large amounts of macro-blocks' matching occur for motion estimation during the processing of a video, gigantic amounts of SAD operations are performed. For example, it requires almost 66,846,720 SAD operations or 534,773,760 SAD operations to process only one single frame of an HD video with the resolution of 1920x1080 or an UHD video with the resolution of 3840x2160, respectively, using a 64x64 macro-block size and a [−64, 63] search range. Furthermore, performing SAD operations on different block sizes shall increase these numbers by multiples. This is hardware-costly and also makes the video processing encounter considerable delay and power consumption. Hardware implementation with parallel SAD operations is the ultimate solution to meet the real-time constraints when using conventional SAD in full search motion estimation.

Bit-truncation has been proposed in [21]-[22] as a promising approach to reduce the cost of the conventional SAD unit. The main concept of bit-truncation is to perform conventional SAD

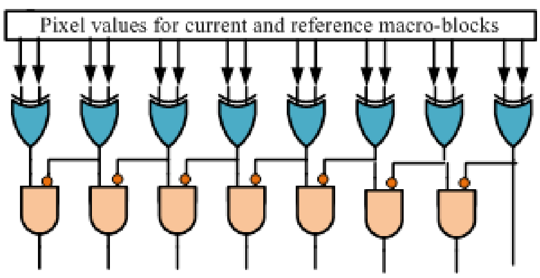


Figure 3. Matching unit without bit-shrinking (MXOR).

The number of generated output bits is 8.

operations on pixels using fewer number of bits by truncating the least significant bits. It is described by:

$$BT_m(u,v) = \sum_{i=0}^{N-1}\sum_{j=0}^{N-1} |p_k\,(i,j)_{7:m} - p_{k-1}\,(i+u,j+v)_{7:m}|; \quad (3)$$

where $p_k\,(i,j)_{7:m}$ is the current pixel of a macro-block with m^{th} least significant bits truncated, $p_{k-1}\,(i+u,j+v)_{7:m}$ is the candidate pixel of a macro-block of a reference frame with m^{th} least significant bits truncated. The macro-block size is $N \times N$, the search range is $[-q, q-1]$ and $-q \le u, v \le q-1$.

The more the number of truncated bits is, the simpler is the SAD operation. As can be deduced from Equation (3), the bit-truncation approach does not really simplify the full adder circuit itself which makes the propagation delay only shortened by the number of truncated bits. Bit-truncation approach also requires a special memory design to support such variable number of bits which are fed into the SAD unit [21].

Considering the hardware-costly SAD and its long processing time and large power consumption, we introduced in [24] a block matching unit with modified XOR functionality that produces almost the same outputs of the SAD unit for the many different combinations of the inputs and is built from a much smaller number of logic gates. It is described by:

$$MXORU(u,v) = \sum_{i=0}^{N-1}\sum_{j=0}^{N-1} \widehat{\grave{XOR}}(p_k\,(i,j), p_{k-1}\,(i+u,j+v)); \quad (4)$$

where,

$$\widehat{XOR} = \sum_{m=0}^{7} (2^m)\left(XOR\left(b_m^{p_k}, b_m^{p_{k-1}}\right)\right).\overline{\left(XOR\left(b_{m-1}^{p_k}, b_{m-1}^{p_{k-1}}\right)\right)}; \quad (5)$$

where $b_m^{p_k}$ is the m^{th} bit of a pixel in the k^{th} frame. We denote the matching unit, described by Equation (5) and shown in Figure 3, by MXOR.

The proposed unit outperforms SAD in terms of hardware-complexity, processing delay and power consumption as explained briefly below.

In a regular Ripple Carry Adder, which is used in the computation of SAD as mentioned before, the highest significant bit of a result depends on the carry which is generated after summing all lower significant bits. This encounters a long delay, since the carry keeps propagating through the adder as explained earlier. In the proposed matching unit, the computation of each output bit depends only on the neighbouring lower significant bit. Thus, the propagation of bits is limited to only one bit position. This saves a considerable amount of time in producing the result of the matching process. Furthermore, the circuit of the proposed matching unit which evaluates each bit is composed of only an XOR gate, an inverter and an AND gate. Hence, the number of constructing logic gates of the proposed unit is enormously less than that of SAD. This is clearly illustrated when observing and comparing the units shown in Figure 2 and Figure 3. The simplification in designing the unit leads to a shorter processing delay along with a smaller unit area and less power consumption.

Since most - and not all - of the output bits of SAD and the modified XOR match, a deterioration in the performance of the video codec shall occur. By implementing the modified

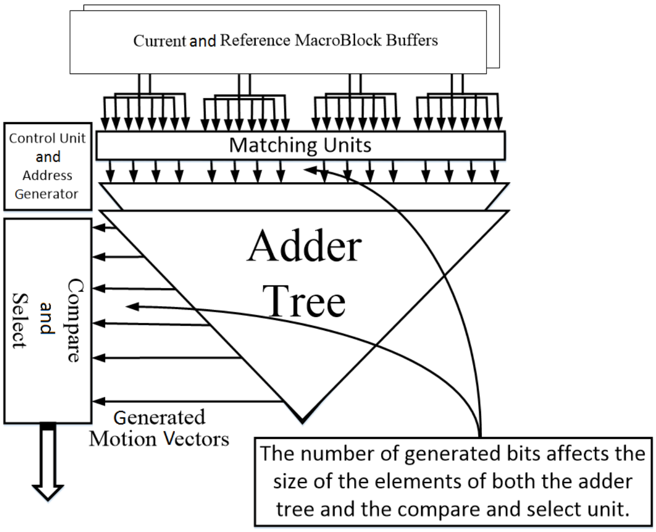


Figure 4. A typical motion estimation engine with a matching unit, an adder tree and a compare and select unit.

XOR in the motion estimation engine of H.264/AVC video codec and taking sample CIF and QCIF videos, we showed in [24] that the amounts of deterioration in both video quality and compression ratio are marginal. In this work, we introduce a new approach, bit-shrinking, to further reduce the computational burden of motion estimation based on the previously introduced MXORU. It is explained as follows.

Bit-shrinking reduces the number of the generated output bits of the matching unit. This is directly reflected on the following stages; namely the parallel adder tree and the compare and select unit shown in Figure 4. The purpose of the adder tree is to accumulate the generated output bits of the matching unit for each sub-block. Therefore, reducing the number of the output bits will obviously reduce the complexity of the processing elements and the used registers in the following stages. Hence, the hardware architecture is simplified and all the related performance metrics in terms of processing delay, power consumption and hardware cost are reduced.

Considering Equation (4) as the general form representation of the block matching unit, the $\widehat{XOR}$ represents several circuit designs according to the level of bit-shrinking. To simplify the Boolean representation of the $\widehat{XOR}$ function of the matching unit, let us first define Ω_m and Ψ_m as:

$$\Omega_m = XOR\left(b_m^{p_k}, b_m^{p_{k-1}}\right), \tag{6}$$

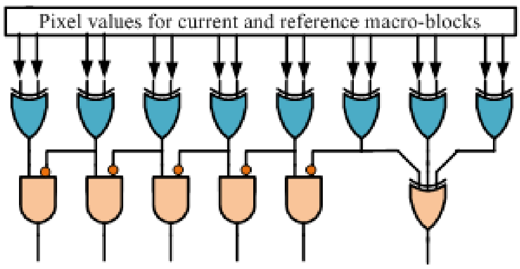


Figure 5. Matching unit with two-bit-shrinking (MXOR2). The number of generated output bits is reduced to 6.

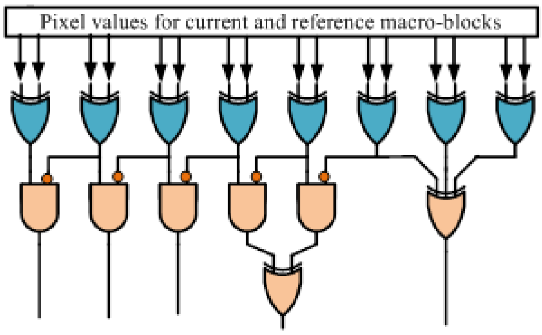


Figure 6. Matching unit with three-bit-shrinking (MXOR3). The number of generated output bits is reduced to 5.

and

$$\Psi_m = \left(XOR\left(b_m^{p_k}, b_m^{p_{k-1}}\right)\right).\overline{\left(XOR\left(b_{m-1}^{p_k}, b_{m-1}^{p_{k-1}}\right)\right)}; \tag{7}$$

where $b_m^{p_k}$ is the m^{th} bit of a pixel in the k^{th} frame.

First, by combining the lowest three output bits of the unit shown in Figure 3 through an XOR gate, one bit is generated. It represents the least significant bit of the output and is described by $2^0 XOR(\Omega_0, \Omega_1, \Omega_2)$. The remaining bits are shifted by 2^{m-2}. Thus, the number of the generated output bits is reduced from eight to six. The matching unit is denoted by MXOR2 and shown in Figure 5. It is also described by:

$$\widehat{XOR} = 2^0 XOR(\Omega_0, \Omega_1, \Omega_2) + \sum_{m=3}^{7} (2^{m-2})\, \Psi_m. \tag{8}$$

Combining the 2nd and 3rd output bits of the unit shown in Figure 5 through another XOR gate reduces the number of the generated output bits to five rather than six.

The least significant bit is described by $2^0 XOR(\Omega_0, \Omega_1, \Omega_2)$. The next higher significant bit, which is generated based on the new combination, is described by $2^1 XOR(\Psi_3, \Psi_4)$ and the remaining output bits are shifted by 2^{m-3}. The matching unit is denoted by MXOR3 and shown in Figure 6. It is also described by:

$$\widehat{XOR} = 2^0 XOR(\Omega_0, \Omega_1, \Omega_2) + 2^1 XOR(\Psi_3, \Psi_4) + \sum_{m=5}^{7} (2^{m-3})\, \Psi_m. \tag{9}$$

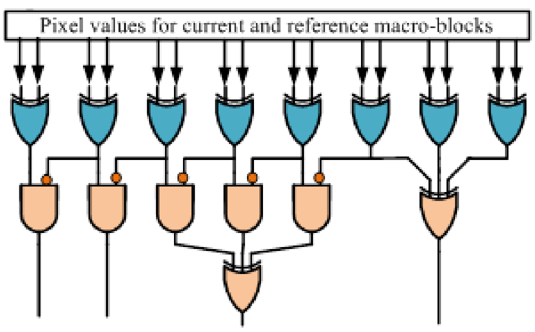


Figure 7. Matching unit with four-bit-shrinking (MXOR4). The number of generated output bits is reduced to 4.

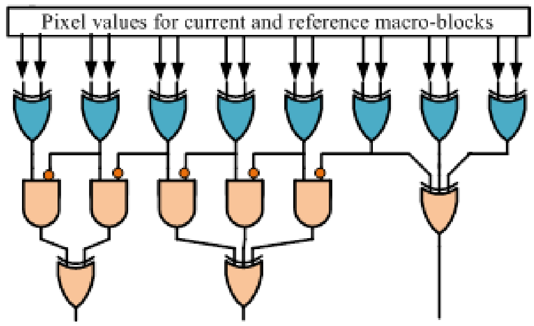


Figure 8. Matching unit with five-bit-shrinking (MXOR5). The number of generated output bits is reduced to 3.

In a different manipulation, combining the 2nd, 3rd and 4th output bits of the unit shown in Figure 5 through an XOR gate reduces the number of the generated output bits to four.

The least significant bit is described by $2^0 XOR(\Omega_0, \Omega_1, \Omega_2)$. The next higher significant bit, which is generated based on the new combination of the three bits, is described by $2^1 XOR(\Psi_3, \Psi_4, \Psi_5)$ and the remaining output bits are shifted by 2^{m-4}. The matching unit is denoted by MXOR4 and shown in Figure 7. It is also described by:

$$\widehat{XOR} = 2^0 XOR(\Omega_0, \Omega_1, \Omega_2) + 2^1 XOR(\Psi_3, \Psi_4, \Psi_5) + \sum_{m=6}^{7} (2^{m-4})\, \Psi_m. \qquad (10)$$

Furthermore, combining the two most significant output bits of the unit shown in Figure 7 through an XOR gate reduces the number of the generated output bits to only three bits.

The least significant bit is described by $2^0 XOR(\Omega_0, \Omega_1, \Omega_2)$. The next higher significant bit is described by $2^1 XOR(\Psi_3, \Psi_4, \Psi_5)$ and finally the most significant bit, which is generated based on the new combination, is described by $2^2 XOR(\Psi_6, \Psi_7)$. The matching unit is denoted by MXOR5 and shown in Figure 8. It is also described by:

$$\widehat{XOR} = 2^0 XOR(\Omega_0, \Omega_1, \Omega_2) + 2^1 XOR(\Psi_3, \Psi_4, \Psi_5) + 2^2 XOR(\Psi_6, \Psi_7). \qquad (11)$$

The generated output bits of the proposed matching units deviate from the original generated output bits of SAD. The more the shrinking is, the more is the deviation. This leads to a degradation in the coding performance of the video codec which is illustrated by deteriorations in both the bit-rate and PSNR. In the following section, we show that the amounts of deterioration are marginal and that our proposed approach actually pays off.

4. Numerical Analysis and Evaluations

First, we perform ASIC system logic synthesis using Synopsys's Design Compiler [ver I-2013.12-SP5-10 using TSMC 90nm general-purpose nominal-threshold-voltage library] for the proposed matching units, the conventional SAD, the modified implementation of SAD [23] (we denote it by MISAD) and finally the matching units of the bit-truncation approach. We consider the latter, since this approach is the closest to our proposed bit-shrinking approach. Taking the number of generated output bits as a reference, our units described by Equations (8), (9), (10) and (11) are analogous to the units which generate six output bits (NTB2), five output bits (NTB3), four output bits (NTB4) and three output bits (NTB5), respectively, of the bit-truncation approach.

We consider the following performance metrics for evaluation purposes: the hardware complexity in terms of the number of two-input NAND gates, the processing delay in terms of the critical path in nano-second and the consumed power in micro-watt. The power estimation is under a global operating voltage of 1.1V, a load capacitance of 0.2 pF, an operating frequency of 100MHz and with a medium confidence level. Table 1 shows the obtained results. From the table, there is a huge increase in performance when using all the proposed matching units when compared with the conventional SAD. The number of two-input NAND gates, consumed power and processing delay of MXOR, which generate the same number of output bits of the conventional SAD, are only 18%, 19% and 14% of the number of two-input NAND gates, consumed power and processing delay, respectively, of the conventional SAD, while their values increase to 24%, 24% and 31% of the conventional SAD, respectively, when considering MXOR5 which generates only 3 output bits. Table 1 also shows that the results of MISAD are very close to the results of the conventional SAD. The number of two-input NAND gates, consumed power and processing delay of MISAD are 96.5%, 98.7% and 90.9% of the number of two-input NAND gates, consumed power and processing delay, respectively, of the conventional SAD, which indicates a very modest improvement of this contemporary work [23]. Table 1 also shows the superiority of our matching units when compared with the matching units of the bit-truncation approach with the same number of generated output bits. Thus, in terms of hardware efficiency illustrated by the adopted three metrics, our proposed matching units outperform the conventional SAD, MISAD and all the matching units of the bit truncation-approach.

The benefits gained by bit-shrinking are not limited to the matching unit only, as we mentioned before, but it also affects later stages of motion estimation engine (shown in Figure 4). Bit-shrinking shall reduce the complexity of the adder tree in terms of a lower-width adder and smaller accumulating registers for each sub-block in 64x64 macro-block. Furthermore, the compare and select unit shall also be reduced in terms of the sizes of the compare units as well as the used registers. For example, the size of each accumulating register reduces from 12 bits, 14 bits, 16 bits, 18 bits and 20 bits for 8-bit pixel matching in a conventional SAD for sub-block sizes of 4x4, 8x8, 16x16, 32x32 and 64x64, respectively, down to only 7 bits, 9 bits, 11 bits, 13 bits and 15 bits, respectively, when using MXOR5.

The significant performance elevation introduced by our units may boost the widespread of H.265/HEVC, especially in modest devices such as smartphones which have storage and power constraints. Nevertheless, we have to show that the compression ratio and video quality are not much affected by the proposed units to support our statement. Therefore, we perform intensive We perform the simulations using HEVC reference software HM16.6 to compare the video quality illustrated by PSNR and the compression ratio illustrated by bit-rate. The PSNR is the

Table 1. Comparison of all matching units in terms of consumed power, number of two-input NAND gates and processing delay.

	SAD	***MXOR***		***MXOR2***		***MXOR3***		***MXOR4***		***MXOR5***	
Number of gates	174	32	(18%)	35	(20%)	37	(21%)	43	(25%)	42	(24%)
Critical path (ns)	20.23	2.87	(14%)	4.2	(21%)	4.42	(22%)	6.18	(31%)	6.18	(31%)
Power (µW)	13.1858	2.5583	(19%)	2.7648	(21%)	2.8554	(22%)	3.0737	(23%)	3.227	(24%)

	SAD	***MISAD***		***NTB2***		***NTB3***		***NTB4***		***NTB5***	
Number of gates	174	168	96.5%	108	(62%)	100	(57%)	86	(49%)	54	(31%)
Critical path (ns)	20.23	18.39	90.9%	18.12	(90%)	17.35	(86%)	16.66	(82%)	13.29	(66%)
Power (µW)	13.1858	13.015	98.7%	10.3675	(79%)	7.2736	(55%)	5.367	(41%)	4.203	(32%)

simulations taking into consideration both the compression ratio and video quality for different kinds of frame sequences of the standard.

most frequently used indicator by the research community to measure picture quality. It is defined by:

$$PSNR = 20\log\frac{255}{\sqrt{MSE}}; \quad (12)$$

where 255 is the largest pixel value for 8-bit representation and MSE is the Mean Square Error between a noise-free M×N monochrome frame I and its noisy approximation K. It is defined by:

$$MSE = \frac{1}{MN}\sum_{i=0}^{M-1}\sum_{j=0}^{N-1}[I\,(i,j) - \,K(i,j)]^2. \quad (13)$$

The simulations are conducted first on six HD videos: Tennis, Beauty, Bosphorus, Honey Bee, Jockey and Ready Steady Go. The resolution of these videos is 1080p. We consider main profile level 6.2 with Random-Access (IBBB frame sequences) and Low-Delay-P (IPPP frame sequences) configurations. The block search range is [−64, 63] with maximum CU size 64 and maximum CU partition depth 4 with full search motion estimation. The number of frames taken in each sequence is 120. The simulations are carried on Windows 8os platform with Intel i7 extreme @ 2.93GHz CPU and 8GB RAM.

Table 2 shows the obtained results of PSNR and the bit-rate of a conventional SAD as well as the deviations in these two metrics when using all matching units compared with SAD. As expected and discussed in the previous section, due to the partial mismatch of the generated output bits between the proposed units and the conventional SAD, a deterioration occurs in both PSNR and bit-rate. Still, the amount of decrease in PSNR and the amount of increase in bit-rate are very small as frankly illustrated by the results obtained from the simulations performed on all the tested videos. Among our proposed matching units, the maximum deviation of PSNR occurs when processing Tennis video with a value of -0.04dB using MXOR5, while the maximum increase in bit-rate also occurs when processing Tennis video with a value of 1.5%. These values are considered very small and hence the deteriorations are actually marginal even at the peaks.

Table 3 summarizes the deviations in PSNR and bit-rate by showing averages. Note that the amount of degradation in the performance is very small when using our approach compared with the conventional SAD. The amount of degradation in PSNR ranges from 0.001dB and 0.002dB when considering MXOR to 0.013dB and 0.018dB when considering MXOR5 for IPPP and IBBB frame sequences, respectively. On the other hand, the amount of increase in the bit-rate ranges from 0.09% and 0.07% when considering MXOR to 0.543% and 0.684% when considering MXOR5 for IPPP and IBBB frame sequences, respectively. It can be deduced from the table that the amounts of deterioration in both PSNR and bit-rate are small when using either bit-truncation or bit-shrinking approaches. Thus, we can confidently state that the video quality and the compression ratio are not practically affected when following these approaches. Note that MISAD performs the SAD operation with a different implementation of the building 1-bit full adder unit as explained earlier in Section 2. Therefore, the generated output bits of MISAD and the conventional SAD are exactly the same. Hence, there are no deteriorations in both bit-rate and PSNR when using MISAD instead of the conventional SAD.

To evaluate the proposed matching units in processing very high resolution videos, we also conduct simulations on 4K UHD videos with the resolution of 3840x2160. Note that UHD videos can be processed by H.265/HEVC and the open-source video codec VP9, but not H.264/AVC. The videos under test are Marathon, Library, Scarf and Traffic & Building. Table 4 summarizes the results by showing averages. Note that the amounts of deviation in both bit-

Table 2. The amounts of deterioration in PSNR (dB) and bit-rate of all units when compared with SAD. The results shown are for both sequences IPPP and IBBB (HD videos).

		SAD		***MXOR***		***MXOR2***		***MXOR3***		***MXOR4***	
VIDEO	***SEQ.***	*BR*	*PSNR*	*Δ-BR%*	*Δ-PSNR*	*Δ-BR%*	*Δ-PSNR*	*Δ-BR%*	*Δ-PSNR*	*Δ-BR%*	*Δ-PSNR*
TENNIS	***IPPP***	153342	37.1282	0.05%	-0.003	-0.38%	-0.018	0.33%	-0.009	0.75%	-0.017
	IBBB	931558	36.5691	0.30%	-0.008	0.06%	-0.030	0.64%	-0.021	1.22%	-0.038
BEAUTY	***IPPP***	439896	38.3652	0.08%	-0.003	-0.06%	-0.002	0.41%	-0.004	0.51%	-0.004
	IBBB	313972	38.0187	0.10%	-0.001	-0.08%	-0.002	0.29%	-0.005	0.49%	-0.007
BOSPHORUS	***IPPP***	397017	38.2873	0.10%	-0.002	0.12%	-0.001	0.00%	-0.013	0.03%	-0.021
	IBBB	316291	38.2566	-0.22%	-0.001	-0.27%	-0.003	-0.17%	-0.012	0.14%	-0.015
HONEYBEE	***IPPP***	98027	39.1785	0.23%	-0.012	0.49%	-0.005	-0.04%	-0.013	0.40%	-0.010
	IBBB	70723	39.7868	-0.07%	0.007	0.16%	-0.006	0.28%	-0.001	0.22%	-0.008
JOCKEY	***IPPP***	389278	39.5202	0.11%	0.014	-0.02%	0.013	0.53%	0.005	0.85%	0.002
	IBBB	305175	39.5179	0.27%	0.000	0.15%	-0.004	0.65%	-0.020	1.08%	-0.019
R.S.G.	***IPPP***	389278	37.3463	-0.01%	-0.001	0.09%	-0.010	0.27%	-0.011	0.77%	-0.013
	IBBB	305175	37.3082	0.06%	-0.008	0.15%	-0.019	0.20%	-0.015	0.89%	-0.011

		MXOR5		***NTB2***		***NTB3***		***NTB4***		***NTB5***	
VIDEO	***SEQ.***	*Δ-BR%*	*Δ-PSNR*	*Δ-BR%*	*Δ-PSNR*	*Δ-BR%*	*Δ-PSNR*	*Δ-BR%*	*Δ-PSNR*	*Δ-BR%*	*Δ-PSNR*
TENNIS	***IPPP***	0.82%	-0.02	-0.03%	0.000	-0.05%	-0.005	0.03%	-0.009	0.36%	-0.018
	IBBB	1.50%	-0.04	0.13%	-0.002	0.14%	-0.006	0.28%	-0.023	1.19%	-0.047
BEAUTY	***IPPP***	0.41%	0.00	0.02%	0.000	-0.01%	0.001	0.25%	-0.005	0.86%	-0.008
	IBBB	0.34%	-0.01	-0.11%	0.000	-0.04%	-0.002	0.27%	-0.007	0.89%	-0.014
BOSPHORUS	***IPPP***	-0.01%	-0.01	0.37%	0.000	0.19%	0.002	0.24%	-0.013	0.11%	-0.020
	IBBB	0.29%	-0.01	-0.20%	0.001	-0.17%	-0.005	-0.07%	-0.002	-0.11%	-0.014
HONEYBEE	***IPPP***	0.32%	-0.02	-0.06%	-0.005	-0.02%	-0.010	0.12%	-0.011	-0.02%	-0.006
	IBBB	-0.06%	0.00	0.20%	0.000	0.52%	0.001	0.46%	-0.001	0.36%	0.008
JOCKEY	***IPPP***	0.90%	-0.01	0.39%	0.008	0.21%	0.014	0.22%	-0.009	0.67%	0.002
	IBBB	1.01%	-0.01	0.09%	-0.004	0.22%	-0.003	0.40%	-0.010	0.77%	-0.019
R.S.G.	***IPPP***	0.83%	-0.02	0.05%	-0.004	0.05%	-0.010	0.02%	-0.011	0.32%	-0.010
	IBBB	1.01%	-0.03	-0.04%	0.003	-0.01%	-0.002	0.09%	-0.004	0.32%	-0.018

Table 3. The average amounts of deterioration in PSNR (dB) and bit-rate of the proposed units and the units of bit-truncation approach when compared with SAD (HD videos).

	MXOR		***MXOR2***		***MXOR3***		***MXOR4***		***MXOR5***	
SEQ.	*Δ-BR%*	*Δ-PSNR*	*Δ-BR%*	*Δ-PSNR*	*Δ-BR%*	*Δ-PSNR*	*Δ-BR%*	*Δ-PSNR*	*Δ-BR%*	*Δ-PSNR*
IPPP	0.09%	-0.001	0.04%	-0.004	0.25%	-0.007	0.55%	-0.010	0.543%	-0.013
IBBB	0.07%	-0.002	0.03%	-0.011	0.31%	-0.012	0.67%	-0.017	0.684%	-0.018

			NTB2		***NTB3***		***NTB4***		***NTB5***	
SEQ.			*Δ-BR%*	*Δ-PSNR*	*Δ-BR%*	*Δ-PSNR*	*Δ-BR%*	*Δ-PSNR*	*Δ-BR%*	*Δ-PSNR*
IPPP			0.122%	-0.0003	0.063%	-0.0014	0.147%	-0.0095	0.384%	-0.010
IBBB			0.012%	-0.0005	0.109%	-0.0029	0.237%	-0.0078	0.568%	-0.018

Jordanian Journal of Computers and Information Technology (JJCIT), Vol. 2, No. 2, August 2016.

rate and PSNR are still very small when using the proposed matching units compared with the conventional SAD. Therefore, we can conclude that the proposed matching units can be utilized in the motion estimation engine of H.265/HEVC to process different resolution videos with marginal effect on the video quality and the compression ratio.

The obtained results of the extensive simulations that are performed on both HD and UHD videos using HM16.6 validate the correctness of the functionalities of the proposed matching units in replacing SAD. The videos have been processed and encoded successfully using H.265/HEVC with marginal deteriorations in the values of both PSNR and bit rate as illustrated in Table 3 and Table 4 and explained earlier. As a validation step for the Verilog code by which we describe the matching units and performed synthesis using Synopsys's Design Compiler as discussed earlier, we adopt a framework which was proposed in [35] that utilizes MATLAB/SIMULINK [36] and ModelSim [37] concurrently as a verification environment for the hardware design of the matching unit of the motion estimation engine. The verification environment co-simulates the hardware design under verification (DUV) along with its software model. It continuously reports mismatches, if existed, for each processed operation at the pixel level along with their occurrence times. In this work, we adjust the verification environment and implement the Verilog codes of the proposed matching units, the conventional SAD, MISAD and the matching units of the bit-truncation approach, one at a time, along with their codes used in HM16.6. The tests are performed on HD and UHD videos which are uploaded to MATLAB and fed to the verification environment. No mismatches are reported. This ensures the validity of the conducted work.

Table 4. The average amounts of deterioration in PSNR (dB) and bit-rate of the proposed units and the units of bit-truncation approach when compared with SAD (UHD videos).

	MXOR		***MXOR2***		***MXOR3***		***MXOR4***		***MXOR5***	
SEQ.	*Δ-BR%*	*Δ-PSNR*	*Δ-BR%*	*Δ-PSNR*	*Δ-BR%*	*Δ-PSNR*	*Δ-BR%*	*Δ-PSNR*	*Δ-BR%*	*Δ-PSNR*
IPPP	0.005%	-0.00165	-0.002%	-0.00182	0.021%	-0.0025	0.251%	-0.0039	0.236%	-0.005
IBBB	0.011%	-0.00053	0.087%	0.000425	0.040%	-0.00318	0.100%	-0.00795	0.155%	-0.008
			NTB2		***NTB3***		***NTB4***		***NTB5***	
SEQ.			*Δ-BR%*	*Δ-PSNR*	*Δ-BR%*	*Δ-PSNR*	*Δ-BR%*	*Δ-PSNR*	*Δ-BR%*	*Δ-PSNR*
IPPP			-0.07%	-0.00145	0.01%	-0.00068	0.06%	-0.00153	0.03%	-0.004
IBBB			-0.03%	-0.00253	0.04%	0.00015	0.02%	-0.0001	0.08%	-0.004

```
uiSum += abs( piOrg[0] - piCur[0] );
```

(a)

```
ABS  = piOrg[0] ^ piCur[0];
ABS1 = ~(ABS << 1) | 1;
MXOR = ABS & ABS1;
uiSum += MXOR;
```

(b)

Figure 9. (a) The code of conventional SAD operation of HM16.6 (b) The code of MXOR implemented in HM16.6.

5. CONCLUSION

We proposed in this paper a hardware-efficient block matching unit for H.265/HEVC motion estimation engine. We followed a bit-shrinking approach with a modified logic functionality to reduce the number of the building two-input NAND gates, consumed power and processing delay of the matching units. We performed extensive simulations to evaluate our design. We considered HD and UHD videos in our evaluations, since their usage have been spreading tremendously in recent years. The results obtained show the superiority of our approach over SAD, MISAD and the bit-truncation approach taking the three adopted performance metrics into consideration. The results show only small amounts of deterioration in video quality and compression ratio when shrinking the number of output bits. The amounts of increase in the number of two-input NAND gates, consumed power and processing delay when using MXOR units with bit-shrinking instead of the pure MXOR unit are small. Nevertheless, the MXOR units with bit-shrinking actually outperform the pure MXOR, since they significantly reduce the hardware complexity of all subsequent stages including the parallel adder tree and the compare and select units of the motion estimation engine.

It is important to clarify that we propose the matching units only for hardware implementations and not for software implementations. The design considers manipulating the pixels at the bit-level which makes the proposed units run efficiently on hardware. As mentioned above, considerable enhancements have been shown when we evaluated the hardware design of the proposed matching units and compared them with the hardware designs of others. Software implementations of the proposed matching units are not efficient, since extra computations are performed by the motion estimation engine when compared with the conventional SAD. The code shown in Figure 9 (a) is the code extracted from HM16.6 that performs one pixel matching using SAD, while the code shown in Figure 9 (b) is the code implemented in HM16.6 to perform the same operation using MXOR which is taken here as an example. The codes to compute MXOR2 through MXOR5 are even longer than the code of MXOR. Hence, extra time is needed by the software to perform the coding process when using our proposed units instead of the conventional SAD. We emphasize the fact that we only performed the software simulations using HM16.6 to validate the functionalities of the matching units and also to measure the amounts of deterioration in both PSNR and bit-rate when using the proposed units instead of SAD.

REFERENCES

[1] Advanced Video Coding, Rec. ITU-T H.264 and ISO/IEC 14496-10, 2014.

[2] High Efficiency Video Coding, Rec. ITU-T H.265 and ISO/IEC 23008-2, 2015.

[3] D. Grois et al., "Performance Comparison of H. 265/MPEG-HEVC, VP9 and H. 264/MPEG-AVC Encoders," Proc. Picture Coding Symp., pp. 394-397, 2013.

[4] VP9, WebM Project's Next-generation Open Video Codec, 2013.

[5] Y. Li, J. Xiao and W. Wu, "Motion Estimation Based on H.264 Video Coding," Proc. 5th Int. Congress Image and Signal Process., pp. 104-108, 2012.

[6] M.U.K. Khan, M. Shafique and J. Henkel, "AMBER: Adaptive Energy Management for On-chip Hybrid Video Memories," Proc. IEEE/ACM Int. Conf. Computer-Aided Design, pp. 405-412, 2013.

[7] R. Li, B. Zeng and M.L. Liou, "A New Three-step Search Algorithm for Block Motion Estimation," IEEE Trans. Circuits Syst. Video Technol., vol. 4, no. 4, pp. 438-442, August 1994.

[8] L. M. Po and W. C. Ma, "A Novel Four Step Search Algorithm for Fast Block Motion Estimation," IEEE Trans. Circuits Syst. Video Technol., vol. 6, no. 3, pp. 313–317, June 1996.

[9] C. Zhu et al., "A Novel Hexagon-based Search Algorithm for Fast Block Motion Estimation," Proc. Int. Conf. Acoustics, Speech and Signal Process., pp. 1593–1596, 2001.

[10] S. Zhu and K.K. Ma, "A New Diamond Search Algorithm for Fast Block-matching Motion Estimation," IEEE Trans. Image Process., vol. 9, no. 2, pp. 287–290, Feburuary 2000.

[11] Y. Nie and K. K. Ma, "Adaptive Rood Pattern Search for Fast Block-matching Motion Estimation," IEEE Trans. Image Process., vol. 11, no. 12, pp. 1442–1448, December 2002.

[12] J. Olivaresa et al., "SAD Computation Based on Online Arithmetic for Motion Estimation," Elsevier Microprocessors and Microsystems, vol. 30, no. 5, pp. 250–258, August 2006.

[13] T. H. Tran, H. M. Cho and S. B. Cho, "Performance Enhancement of Sum of Absolute Difference (SAD) Computation in H.264/AVC Using Saturation Arithmetic," Proc. Emerging Intelligent Computing Technol. and Applicat., pp. 396–404, 2009.

[14] J. Vanne et al., "A High-performance Sum of Absolute Difference Implementation for Motion Estimation," IEEE Trans. Circuits Syst. Video Technol., vol. 16, no. 7, pp. 876-883, July 2006.

[15] S. Lee, J. M. Kim and S. I. Chae, "New Motion Estimation Algorithm Using Adaptively Quantized Low Bit Resolution Image and Its VLSI Architecture for MPEG2 Video Encoding," IEEE Trans. Circuits Syst. Video Technol., vol. 8, no. 6, pp. 734–744, October 1998.

[16] H. Yeo and Y. H. Hu, "A Novel Architecture and Processor-level Design Based on a New Matching Criterion for Video Compression," Proc. IEEE Workshop on VLSI Signal Process. IX, pp. 448–457, 1996.

[17] S. Ertürk, "Multiplication-free One-bit Transform for Low-complexity Block-based Motion Estimation," IEEE Signal Process. Lett., vol. 14, no. 2, pp. 109-112, February 2007.

[18] A. Akin, Y. Dogan and I. Hamzaoglu, "High Performance Hardware Architectures for One Bit Transform Based Motion Estimation," IEEE Trans. Consum. Electron., vol. 55, no. 2, pp. 941–949, August 2009.

[19] A. Akin, G. Sayilar and I. Hamzaoglu, "High Performance Hardware Architectures for One Bit Transform Based Single and Multiple Reference Frame Motion Estimation," IEEE Trans. Consum. Electron., vol. 56, no. 2, pp. 1144–1152, July 2010.

[20] S. Chatterjee and I. Chakrabarti, "Low Power VLSI Architectures for One Bit Transformation Based Fast Motion Estimation," IEEE Trans. Consum. Electron., vol. 56, no.4, pp. 2652–2660, January 2011.

[21] A. Bahari, T. Arslan and A.T. Erdogan, "Low-power H. 264 Video Compression Architectures for Mobile Communication," IEEE Trans. Circuits Syst. Video Technol., vol. 19, no. 9, pp. 1251-1261, September 2009.

[22] I. Chakrabarti, K. Batta and S. Chatterjee, "Efficient Pixel Truncation Algorithm and Architecture," Motion Estimation for Video Coding, Studies in Computational Intelligence, Springer, vol. 590, pp. 65-83, 2015.

[23] D. V. Manjunatha and G. Sainarayanan, "Low-Power Sum of Absolute Difference Architecture for Video Coding," Emerging Research in Electronics, Computer Science and Technol., Lecture Notes in Electrical Engineering, Springer, vol. 248, pp. 335-341, 2014.

[24] E. AlQaralleh and O. M. F. Abu-Sharkh, "Low-complexity Motion Estimation Design Using Modified XOR Function," Springer Multimedia Tools and Applicat., DOI: 10.1007/s11042-015-2948-z, September 2015.

[25] G. Sanchez, M. Porto and L. Agostini, "A Hardware Friendly Motion Estimation Algorithm for the Emergent HEVC Standard and Its Low Power Hardware Design," Proc. 20th IEEE Int. Conf. Image Process., pp. 1991-1994, 2013.

[26] M. E. Sinangil et al., "Hardware-aware Motion Estimation Search Algorithm Development for High-efficiency Video Coding (HEVC) Standard," Proc. 19th IEEE Int. Conf. Image Process., pp. 1529-1532, 2012.

[27] G. Sanchez et al., "DMPDS: A Fast Motion Estimation Algorithm Targeting High Resolution Videos and Its FPGA Implementation," Int. J. Reconfigurable Computing, vol. 2012, pp. 1-12, January 2012.

[28] E. Jaja et al., "Efficient Motion Estimation Algorithms for HEVC/H. 265 Video Coding," Information Science and Applicat., Lecture Notes in Electrical Engineering, Springer, vol. 339, pp. 287-294, 2015.

[29] A. Medhat et al., "Fast Center Search Algorithm with Hardware Implementation for Motion Estimation in HEVC Encoder," Proc. 21st IEEE Int. Conf. Electronics, Circuits and Systems, pp. 155-158, 2014.

[30] K. Miyazawa et al., "Real-time Hardware Implementation of HEVC Video Encoder for 1080p HD Video," Proc. Picture Coding Symp., pp. 225-228, 2013.

[31] G. Pastuszak and M. Trochimiuk, "Algorithm and Architecture Design of the Motion Estimation for the H. 265/HEVC 4K-UHD Encoder," Springer J. Real-Time Image Process., DOI: 10.1007/s11554-015-0516-4, July 2015.

[32] X. Ye, D. Ding and L. Yu, "A Hardware-oriented IME Algorithm and Its Implementation for HEVC," Proc. IEEE Visual Commun. and Image Process. Conf., pp. 205-208, 2014.

[33] Synopsys, (2016, Mar 11), Design Compiler [Online], Available: https://www.synopsys.com/Tools/Implementation/RTLSynthesis/Pages/default.aspx.

[34] JCT-VC High Efficiency Video Coding Reference Software, (2016, Mar 11), HM 16.6 [Online], Available: https:/hevc.hhi.fraunhofer.de.

[35] E. A. AlQaralleh, O. M. F. Abu-Sharkh and B. A. Y. AlQaralleh, "MATLAB/Simulink-based Verification Environment for Motion Estimation in H. 264/AVC," Proc. 5th IEEE Int. Conf. Digital Inform. and Commun. Technol. and Its Applicat., pp. 59-63, 2015.

[36] Mathworks, (2016, Mar 11), Matlab/Simulink Users Guide, Application Program Interface Guide [Online], Available: http://www.mathworks.com.

[37] Mentor Graphics, (2016, Mar 11), ModelSim [Online], Available: https://www.mentor.com/products/fpga/model/

ملخص البحث:

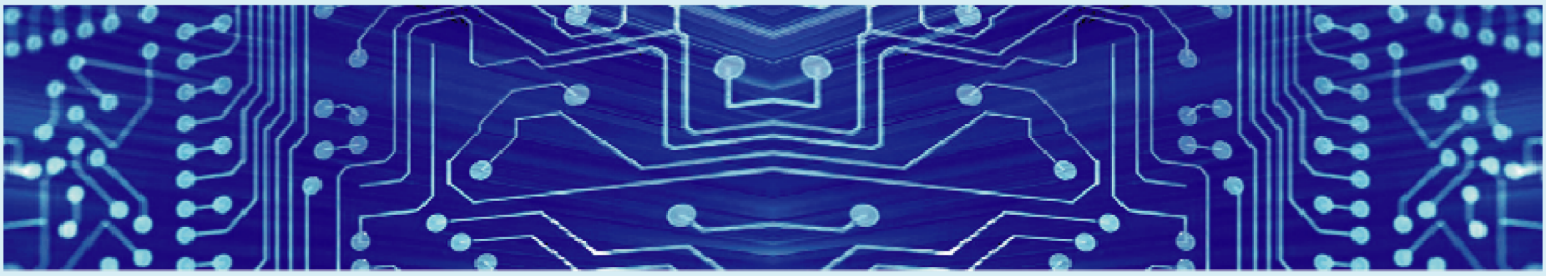
يتمثل الهدف الأساسي من هذا البحث في تحسين المعالجة لِتقنية ترميز الصور عالية الفعالية التي تم إدخالها حديثاً ويرمز لها بـِ (H.265/HEVC). ونظراً لأنّ معظم الحسابات لتلك التقنية ما زالت تُجرى في محرّك تقدير الحركة الموروث من التقنية السابقة المعروفة بـِ (H.264/AVC)، فقد اقترحنا نهجاً يقوم على تقليص عدد خانات الرقم الثنائي بوظيفيّة منطقية معدَّلة من أجل تصميم وحدةٍ لمواءمة الزُّمَر، فعّالة ومبسّطة، لتحلّ محلّ وحدة مجموع الفروق المطلقة المستخدمة سابقاً. وقد تم التقليل من تعقيد الأجزاء المادية للوحدة المقترحة، وكذلك التقليل من عدد ما تقوم بتوليده من خانات الرقم الثنائي الخَرج، الأمر الذي من شأنه أن يبسّط جميع الوحدات اللاحقة لتقدير الحركة.

المجلة الأردنية للحاسوب و تكنولوجيا المعلومات

آب ٢٠١٦ المجلد ٢ العدد ٢
ISSN 2415 - 1076 (Online)
ISSN 2413 - 9351 (Print)

JJCIT

www.jjcit.org
jjcit@psut.edu.jo



مجلة علمية عالمية متخصصة محكمة

تصدر بدعم من صندوق دعم البحث العلمي