Hindawi Publishing Corporation
EURASIP Journal on Wireless Communications and Networking
Volume 2008, Article ID 428397, 12 pages
doi:10.1155/2008/428397

Research Article

Adaptive Transmission of Medical Image and Video Using Scalable Coding and Context-Aware Wireless Medical Networks

Charalampos Doukas and Ilias Maglogiannis

Department of Information and Communication Systems Engineering, School of Sciences, University of the Aegean, 83200 Karlovasi, Samos, Greece

Correspondence should be addressed to Ilias Maglogiannis, imaglo@aegean.gr

Received 15 June 2007; Accepted 25 September 2007

Recommended by Yang Xiao

The aim of this paper is to present a novel platform for advanced transmission of medical image and video, introducing context awareness in telemedicine systems. Proper scalable image and video compression schemes are applied to the content according to environmental properties (i.e., the underlying network status, content type, and the patient status). The transmission of medical images and video for telemedicine purposes is optimized since better content delivery is achieved even in the case of low-bandwidth networks. An evaluation platform has been developed based on scalable wavelet compression with region-of-interest support for images and adaptive H.264 coding for video. Corresponding results of content transmission over wireless networks (i.e., IEEE 802.11e, WiMAX, and UMTS) have proved the effectiveness and efficiency of the platform.

1. INTRODUCTION

A number of telemedicine applications exist nowadays, providing remote medical action systems (e.g., remote surgery systems), patient remote telemonitoring facilities (e.g., homecare of chronic disease patients), and transmission of medical content for remote assessment [1–5]. Such platforms have been proved to be significant tools for the optimization of patient treatment offering better possibilities for managing chronic care, controlling health delivery costs, and increasing quality of life and quality of health services in underserved populations. Collaborative applications that allow for the exchange of medical content (e.g., a patient health record) between medical experts for educational purposes or for assessment assistance are also considered to be of great significance [6–8]. Due to the remote locations of the involved actuators, a network infrastructure (wired and/or wireless) is needed to enable the transmission of the medical data. The majority of the latter data are usually medical images and/or medical video related to the patient. Thus, telemedicine systems cannot always perform in a successful and efficient manner. Issues like large data volumes (e.g., video sequences or high-quality medical images), unnecessary data transmission occurrence, and limited network resources can cause inefficient usage of such systems [9, 10]. In addition, wired and/or wireless network infrastructures often fail to deliver the required quality of service (e.g., bandwidth requirements, minimum delay, and jitter requirements) due to network congestion and/or limited network resources. Appropriate content coding techniques (e.g., video and image compression) have been introduced in order to assess such issues [11–13]; however, the latter are highly associated with specific content type and cannot be applied in general. Additionally, they do not consider the underlying network status for appropriate coding and still cannot resolve the case of unnecessary data transmission.

Scalable coding and context-aware medical networks can overcome the aforementioned issues, through performing appropriate content adaptation. This paper presents an improved patient state and network-aware telemedicine framework. The scope of the framework is to allow for medical image and video transmissions, only when determined to

be necessary, and to encode the transmitted data properly according to the network availability and quality, the user preferences, and the patient status. The framework's architecture is open and does not depend on the monitoring applications used, the underlying networks, or any other issues regarding the telemedicine system used. A prototype evaluation platform has been developed in order to validate the efficiency and the performance of the proposed framework. WiMAX [14], UMTS, and 802.11e network infrastructures have been selected for the networking of the involved entities. The latter wireless technologies provide wide area network connectivity and quality of service (QoS) for specified types of applications. They are considered thus to be suitable for delivering scalable coded medical video services since the QoS classes can be associated with scalable compression schemes. Through the concomitance of the advanced scalable video and image coding and the context-awareness framework, medical video and image delivery can be optimized in terms of better resources utilization and best perceived quality. For example, in the case of patient monitoring, where constant video transmission is required, higher-compression schemes in conjunction with lower QoS network classes might be selected for the majority of content transmission, whereas in case of an emergency event, lower-compression and high QoS classes provide better content delivery for proper assessment. In addition, when a limited resource network is detected (e.g., due to low-bandwidth or high-congestion conditions), video can be replaced by still images transmission. Different compression and transmission schemes may also apply depending on the severity of the case, for example, content transmission for educational purposes versus a case of telesurgery. A scalable wavelet-based compression scheme with region-of-interest (ROI) support [13] has been developed and used for the coding of still medical images, whereas in the case of video, an implementation of scalable H.264 [15] coding has been adopted.

The rest of the paper is organized as follows. Section 2 presents related work in the context of scalable coding and adaptive image and video telemedicine systems. Section 3 describes the proposed scalable image coding scheme, whereas Section 4 deals with the scalable H.264 video coding. Section 5 introduces the proposed context-awareness framework. Performance aspects using a prototype evaluation platform are discussed in Section 6. Finally, Section 7 concludes the article and discusses future work.

2. RELATED WORK IN SCALABLE CODING AND ADAPTIVE IMAGE AND VIDEO TELEMEDICINE SYSTEMS

Scalable image and video coding has attracted recently the interest of several networking research groups from both the academia and the industry since it is the technology that enables the seamless and dynamic adaptation of content to network and terminal characteristics and user requirements. More specifically, scalable coding refers to the creation of a bitstream containing different subsets of the same media (image or video). These subsets consist of a basic layer that provides a basic approximation of the media using an efficient compression scheme and additional datasets, which include additional information of the original image or video increasing the media resolution or decreasing the distortion. The key advantage of scalable coding is that the target bitrate or reconstruction resolution does not need to be known during coding and that the media do not need to be compressed multiple times in order to achieve several bitrates for transmission over various network interfaces. Another key issue is that in scalable coding, the user may determine regions of interest (ROIs) and compress/code them at different resolution or quality levels. This feature is extremely desired in medical images and videos transmitted through telemedicine systems with limited bandwidth since it allows at the same time for zero loss of useful diagnostic information in ROIs and significant compression ratios which result in lower transmission times.

The concept of applying scalable coding in medical images is not quite new. The JPEG2000 imaging standard [16] has been tested in previous published works on medical images [17]. The standard uses the general scaling method which scales (shifts) coefficients so that the bits associated with the ROI are placed in higher bitplanes than the bits associated with the background. Then, during the embedded coding process, the most significant ROI bitplanes are placed in the bitstream before any background bitplanes of the image. The scaling value is computed using the MAXSHIFT method, also defined within the JPEG2000 standard. In this method, the scaling value is computed in such a way that it is possible to have arbitrary shaped ROIs without the need for transmitting shape information to the decoder. The mapping of the ROI from the spatial domain to the wavelet domain is dependent on the used wavelet filters and it is simplified for rectangular and circular regions. The encoder scans the quantized coefficients and chooses a scaling value S such that the minimum coefficient belonging to the ROI is larger than the maximum coefficient of the background (non-ROI area).

A major drawback, however, of the JPEG2000 standard is the fact that it does not support lossy-to-lossless ROI compression. Lossless compression is required in telemedicine systems when the remote diagnosis is based solely on the medical image assessment. In [18], a lossy-to-lossless ROI compression scheme based on set partitioning in hierarchical trees (SPIHTs) [19] and embedded block coding with optimized truncation (EBCOT) [20] is proposed. The input images are segmented into the object of interest and the background, and a chain code-based shape coding scheme [21] is used to code the ROI's shape information. Then, the critically sampled shape-adaptive integer wavelet transforms [22] are performed on the object and background images separately to facilitate lossy-to-lossless coding. Two alternative ROI wavelet-based coding methods with application to digital mammography are proposed by Penedo et al. in [24]. In both methods, after breast region segmentation, the region-based discrete wavelet transform (RBDWT) [23] is applied. Then, in the first method, an object-based extension of the set partitioning in hierarchical trees (OB-SPIHTs) [19] coding algorithm is used, while the second method uses an object-based extension of the set partitioned embedded block (OB-SPECK) [25] coding algorithm. Using

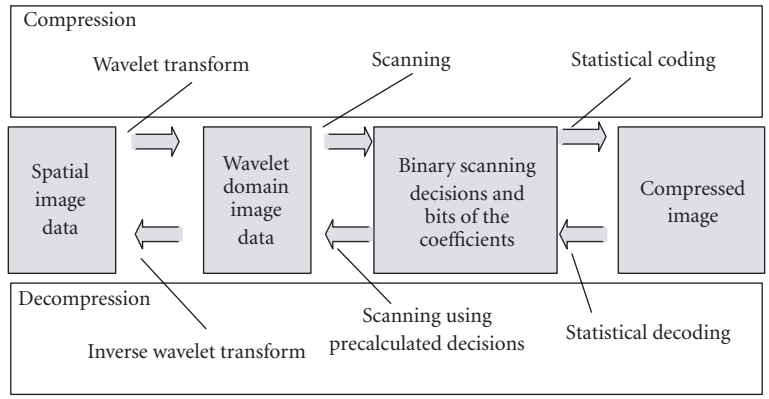

Figure 1: The structure of the DLWIC compression algorithm.

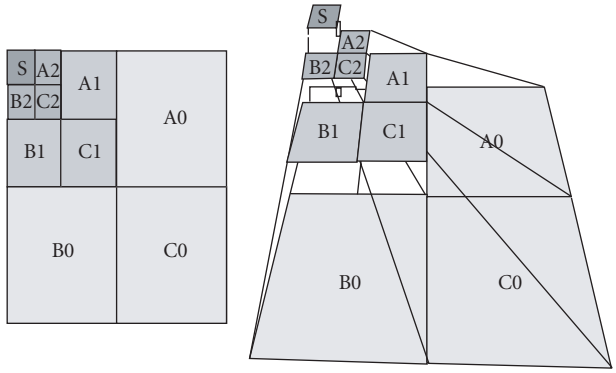

Figure 2: Octave band composition produced by recursive wavelet transform is illustrated on the left and the pyramid structure inside the coefficient matrix is shown on the right.

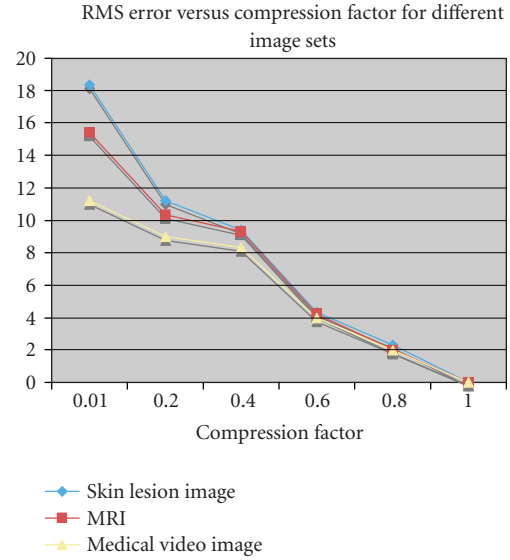

Figure 3: RMS error for different medical images according to quality factors.

RBDWT, it is possible to efficiently perform wavelet subband decomposition of an arbitrary shape region, while maintaining the same number of wavelet coefficients. Both OB-SPIHT and OB-SPECK algorithms are embedded techniques; that is, the coding method produces an embedded bitstream which can be truncated at any point, equivalent to stopping the compression process at a desired quality. The wavelet coefficients that have larger magnitude are those with larger information content. In a comparison, with full-image compression methods as SPIHT and JPEG2000, OB-SPIHT and OB-SPECK exhibited much higher quality in the breast region at the same compression factor [24]. A different approach is presented in [26], where the embedded zerotree wavelets (EZWs) coding technique is adopted for ROI coding in progressive image transmission (PIT). The method uses subband decomposition and image wavelet transform to reduce the correlation in the subimages at different resolutions. Thus, the whole frequency band of the original image is divided into different subbands at different resolutions. The EZW algorithm is applied to the resulting wavelet coefficients to refine and encode the most significant ones.

Scalable video coding (SVC) has been a very active working area in the research community and in international standardizations as well. Video scalability may be handled in different ways: a video can be spatially scalable and can accommodate a range of resolutions according to the network capabilities and the users' viewing screens; it can be temporally scalable and can offer different frame rates (i.e., low frame rates for slow networks); it can be scalable in terms of quality or signal-to-noise ratio (SNR) including different quality levels. In all cases, the available bandwidth of the transmission channel and the user preferences determine resolution, frame rate, and quality of the video sequence. A project on SVC standardization was originally started by the ISO/IEC Moving Picture Experts Group (MPEG). Based on an evaluation of the submitted proposals, the MPEG and the ITU-T Video Coding Experts Group (VCEG) agreed to jointly finalize the SVC project as an amendment of their H.264/MPEG4-AVC standard [15], for which the scalable extension of H.264/MPEG4-AVC, as proposed in [34], was selected as the first working draft. As an important feature of the SVC design, most components of H.264/MPEG4-AVC

are used according to their specification in the standard. This includes the intra- and motion-compensated predictions, the transform and entropy coding, the deblocking, as well as the NAL unit packetization (network abstraction layer (NAL)). The base layer of an SVC bitstream is generally coded in compliance with the H.264/MPEG4-AVC Standard, and each H.264/MPEG-4 AVC Standard-conforming decoder is capable of decoding this base layer representation when it is provided with an SVC bitstream. New tools are only added for supporting spatial and signal-to-noise ratio (SNR) scalability.

Regarding context awareness, despite the numerous implementations and proposals of telemedicine and e-health platforms found in the literature (an indicative reference collection can be found in [1–8]), only a few systems seem to be context-aware. The main goal of context-aware computing is to acquire and utilize information about the context of a device to provide services that are appropriate to particular people, places, times, events, and so forth [40]. According to the latter, the work presented in [41] describes a context-aware mobile system for interhospital communication taking into account the patient's and physician's physical locations for instant and efficient messaging regarding medical events. Bardram presents in [42] additional use cases of context awareness within treatment centers and provides design principles of such systems. The project "AWARENESS" (presented in [43]) provides a more general framework for enhanced telemedicine and telediagnosis services depending on the patient's status and location. To the best of our knowledge, there is no other work exploiting context awareness for optimizing network utilization and efficiency within the context of medical networks and telemedicine services. A more detailed description of the context-aware medical framework is provided in Section 5 along with the proposed implementation.

3. THE PROPOSED SCALABLE IMAGE CODING SCHEME

The proposed methodology adopts the distortion-limited wavelet image codec (DLWIC) algorithm [27]. In DLWIC, the image to be compressed is firstly converted to the wavelet domain using the orthonormal Daubechies wavelet transform [28]. The transformed data are then coded by bitlevels and the output is coded using QM-coder [29], an advanced binary arithmetic coder. The algorithm processes the bits of the wavelet transformed image data in decreasing order concerning their significance in terms of mean square error (MSE). This produces a progressive output stream enabling the algorithm to be stopped at any phase of the coding. The already coded output can be used to construct an approximation of the original image. The latter feature is considered to be useful especially when a user browses medical images using slow bandwidth connections, where the image can be viewed immediately after only few bits have been received; the subsequent bits then make it more accurate. DLWIC uses the progressivism by stopping the coding when the quality of the reconstruction exceeds a threshold given as an input parameter to the algorithm. The presented approach solves the problem of distortion limiting (DL) allowing the user to specify the MSE of the decompressed image. Furthermore, this technique is designed to be as simple as possible consuming less amount of memory in the compression-decompression procedure, thus being suitable for usage on mobile devices.

Figure 1 represents the structure of the DLWIC compression algorithm consisting of three basic steps: (1) the wavelet transform, (2) the scanning of the wavelet coefficients by bitlevels, and (3) the coding of the binary decisions made by the scanning algorithm and the coefficients bits by the entropy encoder. The decoding procedure is almost identical: (1) binary decisions and coefficient bits are decoded; (2) the coefficient data are generated using the same scanning algorithm as in the coding phase, but using the previously coded decision information; (3) the coefficient matrix is converted to a spatial image with the inverse wavelet transform.

The transform is applied recursively to the rows and columns of the matrix representing the original spatial domain image. This operation gives us an octave band composition (see Figure 2). The left side (B) of the resulting coefficient matrix contains horizontal components of the spatial domain image, the vertical components of the image are on the top (A), and the diagonal components are along the diagonal axis (C). Each orientation pyramid is divided into levels; for example, the horizontal orientation pyramid (B) consists of three levels (B0, B1, and B2). Each level contains details of different size; the lowest level (B0), for example, contains the smallest horizontal details of the spatial image. The three orientation pyramids have one shared top level (S), which contains scaling coefficients of the image, representing essentially the average intensity of the corresponding region in the image. Usually, the coefficients in the wavelet transform of a natural image are small on the lower levels and bigger on the upper levels. This property is very important for the compression; the coefficients of this highly skewed distribution can be thus coded using fewer bits.

The coefficient matrix of size $W \times H$ is scanned by bitlevels beginning from the highest bitlevel $n_{\max}$ required for coding the biggest coefficient in the matrix (i.e., the number of the significant bits in the biggest coefficient):

$$n_{\max} = [\log 2(\max\{(|c_{i,j}|)\,|\,0 \le i < W^0 \le j < H\}) + 1], \tag{1}$$

where the coefficient in *(i, j)* is marked with *ci,j*. The coefficients are represented using positive integers as well as the sign bits that are stored separately. The coder first codes all the bits on the bitlevel *nmax* of all coefficients, then all the bits on bitlevel $n_{\max-1}$, and so on until the least significant bitlevel 1 is reached or the scanning algorithm is stopped. The sign is coded together with the most significant bit (the first 1 bit) of a coefficient.

Figure 3 depicts the root mean square (RMS) error results concerning the application of DLWIC algorithm for both lossless (quality factor equal to one) and lossy (quality factor smaller than one) compressions to three different test image sets. The latter consisted of 10 skin lesion images, 10 magnetic resonance images (MRIs), and 10 snapshot images

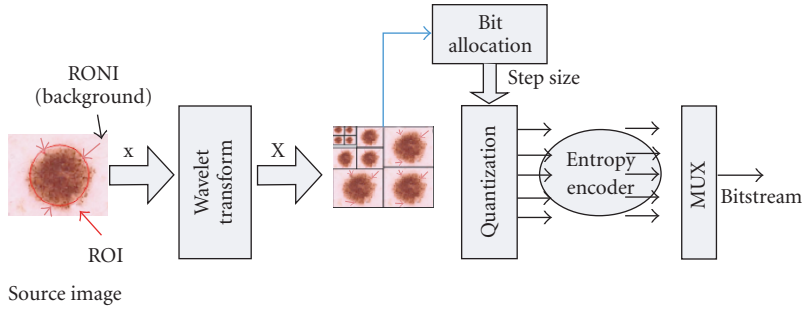

Figure 4: ROI coding system.

Table 1: Average Structural SIMilarity (SSIM) index for three different test image sets using different compression factors. The SSIM index provides an indication of perceptual image similarity between original and compressed images.

		Average SSIM index (%)			
	Compression factor	0.2	0.4	0.6	0.8
Test image	Skin lesion images	86.4475	93.0545	97.0601	99.4466
	MRI images	84.8253	94.1986	97.2828	99.6702
	Medical video snapshots	90.2179	94.5156	97.4221	99.6969

taken from a medical video (see Figure 5 for corresponding images from the aforementioned datasets). With respect to the acquired metrics from the test images, the discussed compression method produces acceptable image quality degradation (RMS value is less than 4 in the case of lossy compression with factor 0.6). For a closer inspection of the compression performance, the Structural SIMilarity (SSIM) index found in [30] is also used as an image quality indicator of the compressed images. The specific metric provides a means of quantifying the perceptual similarity between two images. Perceptual image quality methods are traditionally based on the error difference between a distorted image and a reference image, and they attempt to quantify the error by incorporating a variety of known properties of the human visual system. In the case of SSIM index, the structural information in an image is considered as an attribute for reflecting the structure of objects, independently of the average luminance and contrast, and thus the image quality is assessed based on the degradation of the structural information. A brief literature review [31–33] has shown clearly the advantages of the SSIM index against traditional RMS and peak signal-to-noise ratio (PSNR) metrics and the high adoption by researchers in the field of image and video processing. Average SSIM index values for different compression factors are presented in Table 1. As derived by the conducted similarity comparison experiments using SSIM, the quality degradation even in high compression ratios is not major (i.e., 90.2% and 9.69% for compression factors 0.2 and 0.8, resp., in case of medical video images). This fact proves the efficiency of the proposed algorithm.

In this point, it should be noted that concerning lossy compression, DLWIC performs better in case of medical images of large sizes. Lossy compression is performed by multiplexing a small number of wavelet coefficients (composing the base layer and a few additional layers for enhancement). Thus, a large number of layers are discarded, resulting in statistically higher compression results concerning the file size. However, lossy medical image compression is considered to be unacceptable for performing diagnosis in most of the imaging applications, due to quality degradation that, even minor, can affect the assessment. Therefore, in order to improve the diagnostic value of lossy compressed images, the ROI (region of interest) coding concept is introduced in the proposed application. ROI coding is used to improve the quality in specific regions of interest only by applying lossless or low compression in these regions, maintaining the high compression in regions of noninterest. The wavelet-based ROI coding algorithm implemented in the proposed application is depicted in Figure 4. An octave decomposition is used which repeatedly divides the lower subband into 4 subbands. Let D denote the number of decomposition level, then the number of subbands M is equal to $4+3(D-1)$. Assuming that the ROI shape is given by the client as a binary mask form on the source image, the wavelet coefficients on the ROI and on the region of noninterest (RONI) are quantized with different step sizes. For this purpose, a corresponding binary mask is obtained, called WT mask, on the transform domain. The whole coding procedure can be summarized in the following steps.

(i) The ROI mask is set on the source image x.
(ii) The mask and the requested image x are transferred to the application server.
(iii) The corresponding WT mask B is obtained.
(iv) The DWT coefficient X is calculated.
(v) Bit allocations for the ROI and RONI areas are obtained.
(vi) The X is quantized with the bit allocation from the previous step for each subband of each region.
(vii) The resulting quantized coefficient is encoded.

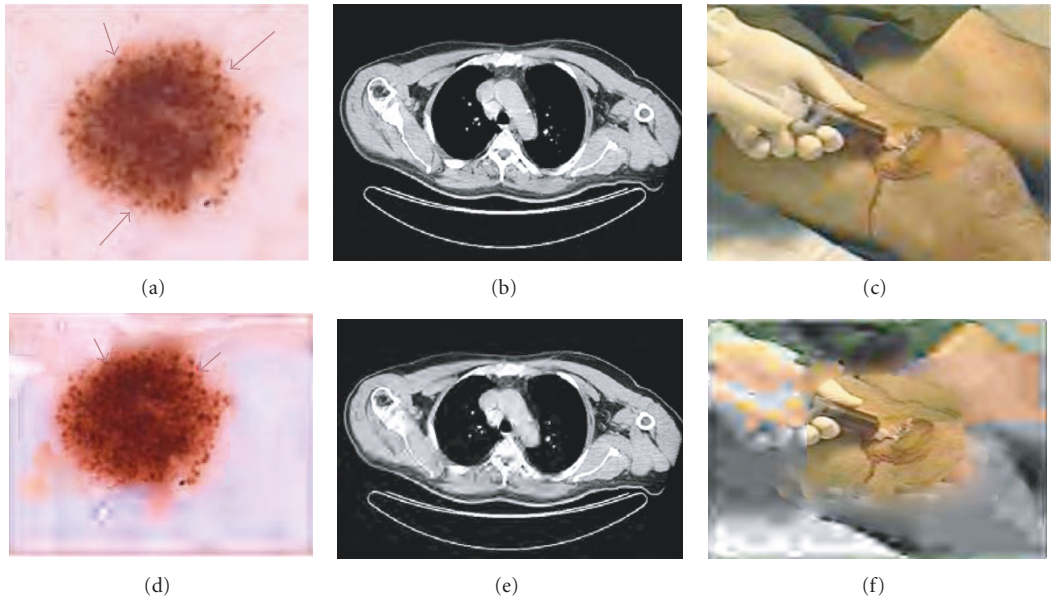

Figure 5: Image samples compressed at different scaling factors and region of interest (ROI) coding. (a) Skin lesion image, (b) MRI image, and (c) medical video image (snapshot) compressed at 0.5 scale factor, respectively. (d)–(f) The same images with background compressed at 0.1 scale factor and ROI at 0.5.

Table 2: Patient data and data levels indicating an urgent status.

Acquired patient data	Data levels indicating an urgent state
ECG (electrocardiogram, 3 leads)	ST wave elevation and depression T-wave inversion
BP (noninvasive blood pressure)	90 mm Hg > systolic > 170 mm Hg
PR (pulse rate)	50/min > PR > 110/min
HR (heart rate)	50/min > HR > 110/min
SpO2 (hemoglobin oxygen saturation)	<90 (%)

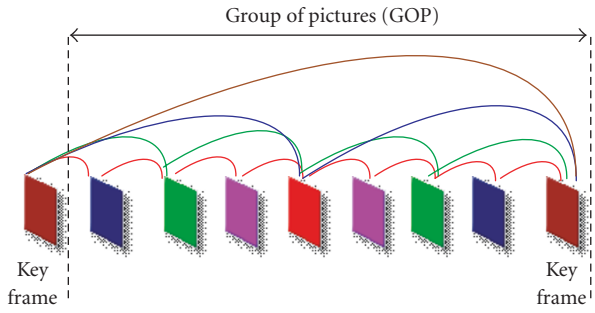

Figure 6: Hierarchical prediction structure of H.264/MPEG-4AVC.

(viii) The WT mask B is encoded.
(ix) The entropy coded coefficient and WT mask are multiplexed in order to create the bitstream.

The decoding process follows the reverse order at the client side. The major advantage of the proposed ROI coding method is that it produces a progressive output stream, and thus the ROI is decoded progressively at the receiver. The user has the capability to stop the transmission at any phase of the coding, while the already transmitted output can be used to construct an approximation of the original image. The specific feature is especially desired for browsing medical images in low-bandwidth mobile networks. In comparison to the JPEG2000 standard, the proposed scheme is preferable since it supports lossy-to-lossless ROI compression. The simplicity of the latter ROI coding requires low computational complexity allowing the usage of the method in real time and even on mobile devices as well.

Figure 5 visualizes the compression effect on image samples from the three different datasets used, using different compression factors and ROI coding. Images (a)–(c) are compressed at 0.5 factor, whereas images (d)–(f) have their background compressed at 0.1 and the ROI at 0.5, respectively. The implementation of the scalable coder has been performed in C++, whereas the encoding part has been developed in Java enabling its usage in both standalone and Web applications (Java applets).

4. H.264 SCALABLE VIDEO CODING

In contrast to older video coding standards as MPEG-2, the coding and display order of pictures are completely decoupled in H.264/MPEG4-AVC standard. Any picture can be marked as reference picture and used for motion-compensated prediction of the following pictures

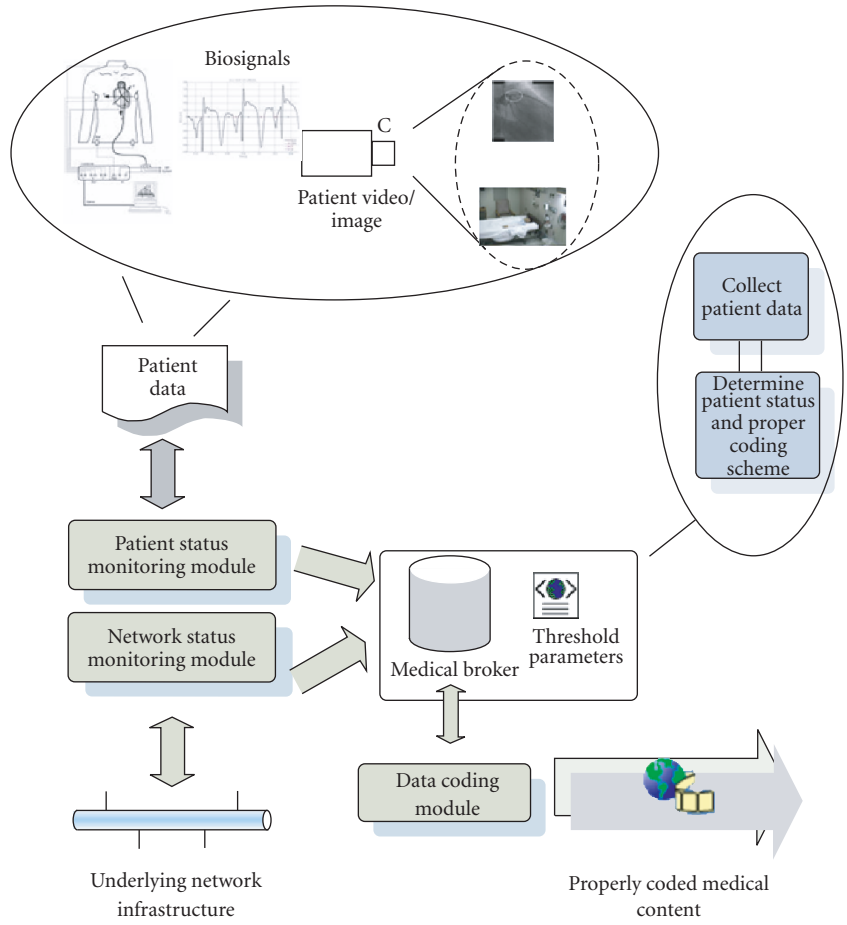

Figure 7: Architecture of the proposed context-aware medical video services framework.

independently of the corresponding slice coding types. These features allow for the coding of picture sequences with arbitrary temporal dependencies.

Temporal scalable bitstreams can be generated by using hierarchical prediction structures as illustrated in Figure 6 without any changes to H.264/MPEG4-AVC. The so-called key pictures are coded in regular intervals by using only previous key pictures as references. The pictures between two key pictures are hierarchically predicted as shown in Figure 6. It is obvious that the sequence of key pictures represents the coarsest supported temporal resolution, which can be refined by adding pictures of the following temporal prediction levels. In addition to enabling temporal scalability, the hierarchical prediction structures also provide an improved coding efficiency compared to classical IBBP coding (named after the corresponding frame sequence) on the cost of an increased encoding-decoding delay [35]. Furthermore, the efficiency of the tools for supporting spatial and SNR scalability is improved as it will be proven in the following sections. It should also be noted that the delay of hierarchical prediction structures can be controlled by restricting the motion-compensated prediction in pictures of the future.

Spatial scalability is achieved by an oversampled pyramid approach. The pictures of different spatial layers are independently coded with layer-specific motion parameters as illustrated in Figure 6. However, in order to improve the coding efficiency of the enhancement layers in comparison to simulcast, additional interlayer prediction mechanisms have been introduced. These prediction mechanisms have been made switchable so that an encoder can freely choose which base layer information should be exploited for an efficient enhancement layer coding. Since the incorporated interlayer prediction concepts include techniques for motion parameter and residual prediction, the temporal prediction structures of the spatial layers should be temporally aligned for an efficient use of the interlayer prediction. It should be noted that all NAL units for a time instant form an access unit and thus have to follow each other inside an SVC bitstream.

5. INTRODUCING THE CONTEXT-AWARENESS FRAMEWORK

This section discusses in detail the proposed context-awareness framework that enables the monitoring of network and patient statuses and determines the appropriate coding of medical video and images using the aforementioned coding techniques. The architecture of the discussed framework is illustrated in Figure 7. The major modules are

Table 3: Video frame, H.264 layer, and WiMAX classes correlation for each scenario.

Video sequences	Frame types			H.264 layers		
	I	P	B	BL	EL1	EL2
A	rtPS	BE	BE	128 Kbps	—	—
B	rtPS	nrtPS	BE	128 Kbps	256 Kbps	256 Kbps

Table 4: Received packet and frame statistics for the evaluation experiments.

Frame type	Packet delay (ms)	Frame delay (ms)	Packet jitter (ms)	Frame jitter (ms)	Packet loss (%)	Frame loss (%)
Video sequence A						
I	302.85	323.45	6.23	7.19	3.2	0.1
P	339.87	322.89	7.14	8.09	12.4	11.1
B	973.86	962.32	9.31	9.43	47.6	47.1
Video sequence B						
I	302.85	323.32	6.71	7.27	3.2	0.1
P	340.81	323.59	7.29	8.17	11.9	11.1
B	942.43	969.36	7.62	8.27	43.7	43.3

(a) the network status monitoring module that determines the current network interface used and the corresponding status, (b) the patient status monitoring module that collects patient data and determines the patient status, and (c) the data coding module which is responsible for properly coding (i.e., compressing) the transmitted video or image, according to instructions given by (d) the medical broker (i.e., usually a repository containing predefined or dynamically defined threshold values for determining patient and network statuses).

The patient state can be determined through a number of biosensors (i.e., heart rate and body temperature sensors) and corresponding vital signals. Defined threshold values in the latter signals determine the case of an immediate video data transmission with better quality (alarm event) to the monitoring unit. In case of normal patient status, periodical video transmission might occur at lower video quality, or alternatively video can be replaced by highly compressed images (suitable for low-bandwidth networks). Video and image coding and transmission can also vary according to network availability and quality. The framework can be also used in cases of remote assessment or telesurgery; according to the network interface used, appropriate video coding is applied to the transmitted medical data, thus avoiding possible transmission delays and optimizing the whole telemedicine procedure. The image and video compression factors are automatically selected based on the current patient status and the underlying network type. Further modification of the latter factors can be performed by the users (i.e., physicians) when the perceived image/video quality is considered to be inappropriate for assessment due to the network conditions.

The framework's architecture is open and does not depend on the monitoring applications used, the underlying networks, or any other issues regarding the telemedicine system used. For this purpose, Web services [36, 37] have been used as a communication mechanism between the major framework components and the external patient monitoring applications used. The message exchange has been implemented through SOAP [38], a simple yet very effective and flexible XML-based communication mechanism. The latter involves the session initialization (which more precisely includes user authentication and service discovery) and the exchange of status and control messages. The status messages include information regarding the patient data as generated from the monitoring sensors and the underlying network status and quality, whereas the control messages contain instructions regarding the proper coding of the transmitted data (see Figure 9). It should be noted that the involved modules for the aforementioned communication (see Figure 8) can all reside at the patient's site, or alternatively the medical broker can reside at the remote treatment site for the direct collection of medical data and the reactive instruction's provision.

The following section provides information regarding the evaluation of the proposed platform using H.264 and wavelet scalable coding for image and video data.

6. EVALUATION PLATFORM

In order to validate the adaptive transmission of medical video and image data using context-aware medical networks, an evaluation platform has been implemented based on the concept described in Section 5. H.264 [15, 34] has been used for video coding and scalable wavelet for image coding [13], respectively. The main components of which the platform consists are the attached biosensors to the patient, the software modules responsible for collecting the corresponding signals and determining the appropriate video coding depending on the patient and network statuses, and the simulated network infrastructures (i.e., IEEE 802.11g (WLAN), UMTS, and WiMAX) for data transmission to the monitoring units (e.g., a treatment center, an ambulance, or a physician at a remote site). Two patient states have been defined: normal and urgent. The patient data that are monitored through corresponding sensors are ECG,

Table 5: Response time for UMTS and WLAN radio segments.

Compression scheme		Response time T_R (s) WLAN | UMTS No compression	JPEG	Wavelet (lossless)	Wavelet (lossy)
Image type (file size)	Skin lesion (520 Kb)	8.6 | 35.2	8.1 | 27	3.6 | 20	3 | 19.6
	MRI (525 Kb)	8.8 | 41.3	8.3 | 28	4.3 | 21.3	3.9 | 19.9
	Video snapshot (1 Mb)	14.2 | 54	8.4 | 29.3	5.1 | 22	4 | 17

Table 6: ROI transmission time for UMTS, WLAN, and WiMAX emulated radio segments.

ROI coding		ROI transmission time (s) UMTS	WLAN	WiMAX
Image type (file size)	CR (262 Kb)	5	1.5	1.42
	CT (525 Kb)	5.7	1.55	1.5
	MR (1 Mb)	6.2	2.1	1.9

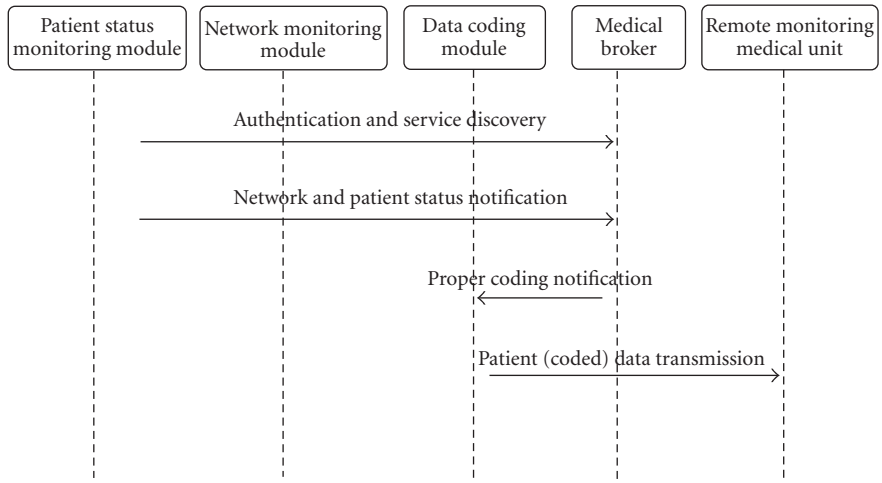

Figure 8: Message exchange between the framework's modules for the case of remote patient status monitoring.

```
<?xml version="1.0" encoding="UTF-8"?>
   <soapenv:Envelope xmlns:soapenv="http://schemas.xmlsoap.org/soap/envelope/"
xmlns:xsd="http://www.w3.org/2001/XMLSchema" xmlns:xsi="http://www.w3.org/2001/
XMLSchema-instance">
      <soapenv:Header>
      </soapenv:Header>
      <soapenv:Body>
         <create xmlns="urn:enterprise.soap.sforce.com">
            <sObjects xsi:type="ns3:EventData"
xmlns:ns3="urn:sobject.enterprise.soap.sforce.com">
               <ns3:EventType>Urgent</ns3:EventType>
               <ns3:Device>PDA</ns3:Device>
               <ns3:NetworkType>GPRS</ns3:NetworkType>
               <ns3:NetworkQuality>Medium</ns3:NetworkQuality>
               <ns3:PatientID>13001</ns3:PatientID>
               <ns3:BloodPressure>75mmHg</ns3:BloodPressure>
               <ns3:PulseRate>55min</ns3:PulseRate>
               <ns3:HeartRate>45/min</ns3:HeartRate>
            </sObjects>
         </create>
      </soapenv:Body>
   </soapenv:Envelope>
```

Figure 9: XML instance of an SOAP message containing information about the patient status.

BP (noninvasive blood pressure), PR (pulse rate), HR (heart rate), and SpO2 (hemoglobin oxygen saturation).

6.1. Evaluation results from H.264 video compression

For[RS10] the evaluation of H.264 video coding, two video sequences have been used for transmission corresponding to the two defined patient statuses. The media access control (MAC) layer of 802.16 enables the differentiation among traffic categories with different multimedia requirements. The standard [44] supports four quality-of-service scheduling types: (1) unsolicited grant service (UGS) for the constant bitrate (CBR) service, (2) real-time polling service (rtPS) for the variable bitrate (VBR) service, (3) nonreal-time polling service (nrtPS) for nonreal-time VBR, and (4) best effort (BE) service for service with no rate or delay requirements. For the specific scenario, a simulated WiMAX wireless network of 1 Mbps has been used; the following rates for the supported traffic classes have been allocated: 200 Kbps for the UGS class, 300 Kbps for the rtPS class, 200 Kbps for the nrtPS classes, and 200 Kbps for the best BE class. Each group of pictures (GOP) is consisted of I, P, and B frames structured by repeating sequences of the period IBBPBBPBB. The GOP contains 25 frames per second, and the maximum UDP packet size is at 1000 bytes (payload only). The NS2 simulator [45] and a WiMAX module presented in [46] have been used for this purpose. A number of 11 nodes randomly distributed at a surface of 1000 m^2 using omnidirectional antenna models provided by NS2 have been simulated.

A scalable extension of H.264/AVC encoder and decoder was used, provided by [39]. A number of background flows are also transmitted in the simulated network in order to fill in the respective WiMAX class capacity in the link. The background traffic is increased from 210 Kbps to 768 Kbps leading the system in congestion. For evaluation purposes, we adopt a simpler QoS mapping policy, by using direct mapping of packets to WiMAX classes. All packets are formed into three groups, according to the type of context that they contain, and each group of packets is mapped to one WiMAX class.

The first simulation scenario refers to the normal patient status. The corresponding video sequence has been used with a single layer H.264 transmission; rtPS for transmitting I frames and nrtPS and BE for transmitting P and B frames are used, respectively. The second simulation scenario refers to the urgent patient status and it considers a scalable H.264 stream transmission consisting of two layers; the base layer (BL) packets are encoded using the scalable extension H.264/AVC codec at 128 kbps and two enhancement layers (ELs) (i.e., EL1 and EL2) are encoded each at 256 kbps, respectively. The correlation between the video frames, the H.264 layers, and the network classes is presented in Table 3.

The experimental results prove better network resources utilization in case of the normal patient status, and acceptable video quality in case of the urgent patient status. PSNR and Structural SIMilarity (SSIM) index [31] have been used as quality metrics. In the case of normal patient status (video sequence A, higher compression, and low network quality class used), PSNR and SSIM were calculated at the receiver at 23.456 and 0.762, respectively. In the case of the urgent patient status (video sequence B, better video coding, and higher network quality class used), PSNR and SSIM were calculated at 29.012 and 0.918, respectively.

Table 4 presents corresponding results regarding the packet and frame statistics at the receiver side for each frame type (i.e., I, P, and B). There is a decrease in frame delay, loss, and jitter for the second video sequence despite the fact that video is encoded in higher quality. The latter is translated into both better network resource utilization and proper video quality when context awareness indicates the proper video coding and transmission schemes.

6.2. Evaluation results from wavelet image compression

Regarding the wavelet image coding, the first set of measurements concerns the framework's response time (i.e., the time to transmit a compressed image) for different image types (skin lesion, MRI, and video snapshot with sizes of 520, 525 Kb, and 1.3 Mb, resp.) for different types of compression (no compression, JPEG compression with a quality factor of 0.75, and lossless and lossy discrete wavelet compression), for the cases where either UMTS or WLAN is the access network. The corresponding results are depicted in Table 5.

The lossless compression can be selected for cases where the underlying network infrastructure has the means (i.e., high bandwidth, limited jitter, and delay) to support transmission of larger data size or in cases where the context of the transmission demands high perceived image quality (e.g., a patient emergency event). In correspondence to the latter, lossy image compression can be used in the case of patient monitoring through still images using a resource-limited wireless network (e.g., UMTS).

With respect to the evaluation results, discrete wavelet compression reduces the actual medical image downloading time improving in this way the response time for the proposed application. An additional performance metric of the proposed medical application concerns the ROI transmission time for the same image dataset for three emulated radio access networks (i.e., UMTS, WLAN, and WiMAX). The corresponding results are depicted in Table 6.

7. CONCLUSION

Medical video and image transmission is a key issue for the successful deployment and usage of telemedicine applications especially when wireless network infrastructures are used. Adaptive and scalable coding on the other hand is considered to be quite important since it is the technology that enables the seamless and dynamic adaptation of content according to network or patient status and their requirements. This paper introduces the concept of adaptive transmission of medical video and image data in telemedicine systems using context-aware medical networks. Adaptive transmission is achieved through scalable video coding using H.264 and wavelet-based scalable image compression with ROI coding support. The simplicity of the latter ROI coding requires

low computational complexity allowing for the usage of the method in real time and even on mobile devices as well.

Context awareness is achieved through the monitoring of the patient status, the context of the data transmission, and the network status. Evaluation results using different wireless networks (i.e., IEEE 802.11e, WiMAX, and UMTS) indicate the effectiveness of the platform in the context of both efficient data compression with acceptable quality degradation and proper data transmission over wireless networks. What remains as future work is the establishment of the innovative context-aware medical video and image coding platform into a real patient-care environment, providing helpful information regarding the assessment of the platform in use.

REFERENCES

[1] J. C. Lin, "Applying telecommunication technology to healthcare delivery," *IEEE Engineering in Medicine and Biology Magazine*, vol. 18, no. 4, pp. 28–31, 1999.

[2] S. Pavlopoulos, E. Kyriacou, A. Berler, S. Dembeyiotis, and D. Koutsouris, "A novel emergency telemedicine system based on wireless communication technology-AMBULANCE," *IEEE Transactions on Information Technology in Biomedicine*, vol. 2, no. 4, pp. 261–267, 1998.

[3] S. Deb, S. Ghoshal, V. N. Malepati, and D. L. Kleinman, "Tele-diagnosis: remote monitoring of large-scale systems," in *Proceedings of IEEE Aerospace Conference*, vol. 6, pp. 31–42, Big Sky, Mo, USA, March 2000.

[4] Y. B. Choi, J. S. Krause, H. Seo, K. E. Capitan, and K. Chung, "Telemedicine in the USA: standardization through information management and technical applications," *IEEE Communications Magazine*, vol. 44, no. 4, pp. 41–48, 2006.

[5] C. S. Pattichis, E. Kyriacou, S. Voskarides, M. S. Pattichis, R. Istepanian, and C. N. Schizas, "Wireless telemedicine systems: an overview," *IEEE Antennas and Propagation Magazine*, vol. 44, no. 2, pp. 143–153, 2002.

[6] M. Akay, I. Marsic, A. Medl, and G. Bu, "A system for medical consultation and education using multimodal human/machine communication," *IEEE Transactions on Information Technology in Biomedicine*, vol. 2, no. 4, pp. 282–291, 1998.

[7] J. Zhou, X. Shen, and N. D. Georganas, "Haptic tele-surgery simulation," in *Proceedings of the 3rd IEEE International Workshop on Haptic, Audio and Visual Environments and their Applications (HAVE '04)*, pp. 99–104, Ottawa, Canada, October 2004.

[8] P. Fontelo, E. DiNino, K. Johansen, A. Khan, and M. Ackerman, "Virtual microscopy: potential applications in medical education and telemedicine in countries with developing economies," in *Proceedings of the 38th Annual Hawaii International Conference on System Sciences (HICSS '05)*, p. 153, Big Island, Hawaii, USA, January 2005.

[9] A.-L. Lage, J. Martins, J. Oliveira, and W. Cunha, "A quality of service approach for managing tele-medicine multimedia applications requirements," in *Proceedings of the 4th IEEE International Workshop on IP Operations and Management (IPOM '04)*, pp. 186–190, Beijing, China, October 2004.

[10] C. LeRouge, M. J. Garfield, and A. R. Hevner, "Quality attributes in telemedicine video conferencing," in *Proceedings of the 35th Annual Hawaii International Conference on System Sciences (HICSS '02)*, pp. 2050–2059, Big Island, Hawaii, USA, January 2002.

[11] H. Yu, Z. Lin, and F. Pan, "Applications and improvement of H.264 in medical video compression," *IEEE Transactions on Circuits and Systems*, vol. 52, no. 12, pp. 2707–2716, 2005.

[12] G. Bernabe, J. Gonzalez, J. M. Garcia, and J. Duato, "A new lossy 3-D wavelet transform for high-quality compression of medical video," in *Proceedings of the IEEE EMBS International Conference on Information Technology Applications in Biomedicine (ITAB '00)*, pp. 226–231, Arlington, Va, USA, November 2000.

[13] C. N. Doukas, I. Maglogiannis, and G. Kormentzas, "Medical image compression using wavelet transform on mobile devices with ROI coding support," in *Proceedings of the 27th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC '05)*, vol. 7, pp. 3779–3784, Shanghai, China, September 2005.

[14] IEEE 802.16 WG, "IEEE Standard for Local and Metropolitan Area Networks—Part 16: Air Interface for Fixed Broadband Wireless Access Systems," IEEE Std. 802.16-2004, October 2004.

[15] ITU-T Rec. & ISO/IEC 14496-10 AVC, "Advanced Video Coding for Generic Audiovisual Services," version 3, 2005.

[16] ISO/IEC JTC 1/SC 29/WG 1 (ITU-T SG8), "JPEG 2000 Final Committee Version 1.0," March 2000.

[17] G. Anastassopoulos and A. Skodras, "JPEG2000 ROI coding in medical imaging applications," in *Proceedings of the 2nd IASTED International Conference on Visualisation, Imaging and Image Processing (VIIP '02)*, pp. 783–788, Palma de Mallorca, Spain, August 2002.

[18] Z. Liu, J. Ha, Z. Xiong, Q. Wu, and K. Castleman, "Lossy-to-lossless ROI coding of chromosome images using modified SPIHT and EBCOT," in *Proceedings of IEEE International Symposium on Biomedical Imaging*, pp. 317–320, Washington, DC, USA, July 2002.

[19] A. Said and W. A. Pearlman, "A new, fast, and efficient image codec based on set partitioning in hierarchical trees," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 6, no. 3, pp. 243–250, 1996.

[20] D. Taubman, "High performance scalable image compression with EBCOT," *IEEE Transactions on Image Processing*, vol. 9, no. 7, pp. 1158–1170, 2000.

[21] Z. Liu, Z. Xiong, Q. Wu, Y.-P. Wang, and K. Castleman, "Cascaded differential and wavelet compression of chromosome images," *IEEE Transactions on Biomedical Engineering*, vol. 49, no. 4, pp. 372–383, 2002.

[22] G. Minami, Z. Xiong, A. Wang, and S. Mehrotra, "3-D wavelet coding of video with arbitrary regions of support," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 11, no. 9, pp. 1063–1068, 2001.

[23] S. Li and W. Li, "Shape-adaptive discrete wavelet transforms for arbitrarily shapedvisual object coding," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 10, no. 5, pp. 725–743, 2000.

[24] M. Penedo, W. A. Pearlman, P. G. Tahoces, M. Souto, and J. J. Vidal, "Region-Based Wavelet Coding Methods for Digital Mammography," *IEEE Transactions on Medical Imaging*, vol. 22, no. 10, pp. 1288–1296, 2003.

[25] A. Islam and W. A. Pearlman, "Embedded and efficient low-complexity hierarchical image coder," in *Visual Communications and Image Processing*, vol. 3653 of *Proceedings of SPIE*, pp. 294–305, San Jose, Calif, USA, January 1999.

[26] R. S. Dilmaghani, A. Ahmadian, M. Ghavami, M. Oghabian, and H. Aghvami, "Multi rate/resolution control in progressive medical image transmission for the region of interest (ROI) using EZW," in *Proceedings of the 25th Annual International*

Conference of the IEEE Engineering in Medicine and Biology Society (EMBS '03), vol. 1, pp. 818–820, Cancun, Mexico, September 2003.

[27] J. M. Shapiro, "Embedded image coding using zerotrees of wavelet coefficients," *IEEE Transactions on Signal Processing*, vol. 41, no. 12, pp. 3445–3462, 1993.

[28] A. Said and W. A. Pearlman, "A new, fast, and efficient image codec based on set partitioning in hierarchical trees," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 6, no. 3, pp. 243–250, 1996.

[29] J. Lehtinen, "Limiting distortion of a wavelet image codec," *Acta Cybernetica*, vol. 14, no. 2, pp. 341–354, 1999.

[30] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, 2004.

[31] Y.-B. Tong, Q. Chang, and Q.-S. Zhang, "Image quality assessing by using NN and SVM," in *Proceedings of the 5th International Conference on Machine Learning and Cybernetics (ICMLC '06)*, vol. 2006, pp. 3987–3990, Dalian, China, August 2006.

[32] G.-H. Chen, C.-L. Yang, L.-M. Po, and S.-L. Xie, "Edge-based structural similarity for image quality assessment," in *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP '06)*, vol. 2, pp. II933–II936, Toulouse, France, May 2006.

[33] Z.-Y. Mai, C.-L. Yang, and S.-L. Xie, "Improved best prediction mode(s) selection methods based on structural similarity in H.264 I-frame encoder," in *Proceedings of IEEE International Conference on Systems, Man and Cybernetics (SMC '05)*, vol. 3, pp. 2673–2678, Waikoloa, Hawaii, USA, October 2005.

[34] H. Schwarz, et al., "Technical Description of the HHI proposal for SVC CE1," ISO/IEC JTC1/WG11, Doc. m11244, Palma de Mallorca, Spain, October 2004.

[35] H. Schwarz, D. Marpe, and T. Wiegand, "Hierarchical B pictures," Joint Video Team, Doc. JVT-P014, Poznan, Poland, July 2005.

[36] M. Hori and M. Ohashi, "Applying XML Web services into health care management," in *Proceedings of the 38th Annual Hawaii International Conference on System Sciences (HICSS '05)*, p. 155, Big Island, Hawaii, USA, January 2005.

[37] Y. Lee, C. Patel, S. A. Chun, and J. Geller, "Towards intelligent Web services for automating medical service composition," in *Proceedings of IEEE International Conference on Web Services*, pp. 384–391, 2004.

[38] SOAP specifications, http://www.w3.org/TR/soap/.

[39] JSVM 0 software, http://ip.hhi.de/imagecom_G1/savce/index.htm.

[40] T. P. Moran and P. Dourish, "Introduction to this special issue on context-aware computing," *Human-Computer Interaction*, vol. 16, no. 2–4, pp. 87–95, 2001.

[41] M. A. Munoz, M. Rodriguez, J. Favela, A. I. Martinez-Garcia, and V. M. Gonzalez, "Context-aware mobile communication in hospitals," *Computer*, vol. 36, no. 9, pp. 38–46, 2003.

[42] J. E. Bardram, "Applications of context-aware computing in hospital work: examples and design principles," in *Proceedings of the ACM Symposium on Applied Computing (SAC '04)*, vol. 2, pp. 1574–1579, Nicosia, Cyprus, March 2004.

[43] T. Broens, A. van Halteren, M. van Sinderen, and K. Wac, "Towards an application framework for context-aware m-health applications," in *Proceedings of the 11th Open European Summer School (EUNICE '05)*, Madrid, Spain, July 2005.

[44] IEEE 802.16 Working Group, "IEEE Standard for Local and Metropolitan Area Networks—Part 16: Air Interface for Fixed Broadband Wireless Access Systems," IEEE Std. 802.16-2004, October 2004.

[45] The Network Simulator—ns2, http://www.isi.edu/nsnam/ns/.

[46] J. Chen, C.-C. Wang, F. C.-D. Tsai, et al., "Simulating wireless networks: the design and implementation of WiMAX module for ns-2 simulator," in *Proceedings of ACM Workshop on ns-2: The IP Network Simulator (WNS2 '06)*, Pisa, Italy, October 2006.