Abstract—We analyze viewer-perceived quality of a compressed video stream, transmitted over a lossy IP network with a quality of service mechanism. The parameters of the encoding schemes include the transmission bit rate, the compression depth, the frame size and the frame rate. We demonstrate that when jointly considering the impact of the coding bit rate, the packet loss ratio and the video characteristics, we can identify an optimal encoding scheme that maximizes viewer-perceived quality. The video content, the compression and the transmission are represented by a vector \( \mathbf{X} \) which contains \( d \) parameters. Based on subjective tests, we obtain a set of observation pairs of labeled samples \( P_L = \{ X_1 \}, Q_1 \}, \ldots \} \), where \( Q_1 \) is the quality class related to the vector of input parameters \( X_1 \). To determine the significance of these results, we use the analysis of variance (ANOVA) statistical method, which identifies those factors that cause differences in the averages in the subjective tests results, and determines the significance of the results. Finally, we introduce a novel method to predict an optimal encoding scheme based on canonical discriminant analysis (CDA) for feature classification.

Index Terms—Encoding scheme, MPEG, perceived quality, QoE, subjective tests, video QoS, video quality.

I. INTRODUCTION

DIGITAL video communication systems over Internet Protocol (IP) networks exhibit diverse types of distortions that are strongly influenced by video characteristics, compression methods and transmission capabilities. Assuring quality of experience (QoE) for IP television is rapidly becoming a top priority for vendors and service providers as the market is evolving dramatically and services are now commercially deployed. QoE models and network quality of service (QoS) are crucial for delivering commercial video services.

Since perceived visual information is affected by the network differently than data or voice services [1], measuring and modeling network capabilities in relation to visual information is imperative for introducing video services over IP networks.

The convergence of video, voice and data into triple play service bundles, running on the same network, is a key strategy for many service providers seeking to establish competitive advantages over the residential consumer market in the forthcoming years. QoS involves the ability of a network element to achieve some level of assurance that its traffic and service requirements can be satisfied [2]. QoS protocols provide the mechanism to differentiate traffic, while policy defines how these protocols are used. Furthermore, service-driven QoS provides enhanced QoS based on the content or application. Video QoE is an example of service-driven QoS system. QoE is a subjective measure of performance in a system, which relies on human opinion, and thus differs from QoS which can be precisely measured.

Video QoE is composed of the quality of the network delivery system, the quality of the encoding/decoding components of the network, and human factors. It is affected by the following factors:

A. Compression Impairments

A lossy compression, such as MPEG, exploits two types of redundancy: spatio-temporal redundancy and psycho-visual redundancy. Examples of compression-related impairments include edge busyness, error blocks, localized smearing, and blocking/tiling [2], [3]. The characteristics of the source video play a crucial role in determining the amount of compression that is possible, and hence, the severity of the compression artifacts [3].

B. Rate Control

Rate control means jointly adjusting some coding parameters in order to achieve a target bit rate for the video sequence transmission [4]–[7]. The compression parameter of the rate control in compression standards is usually quantization. But two additional parameters can be used for rate control—spatial resolution and frame rate. To evaluate the influence of rate control on the video stream quality, subjective tests are often useful.

C. Transmission Impairments

Video transmission over IP networks is considered as a real-time application. In our simulations the video streams are transmitted through networks using UDP/IP at a constant bit rate (CBR), since that best suits the QoS enforcement mechanism [8]. The most important parameters of transmission networks are one-way delay (delay), instantaneous packet delay variation.
(jitter), packet loss rate (PLR), the loss distribution, throughput and available bandwidth [3], [7], [9]–[11]. Since MPEG compression creates a dependency within the sequence, data loss that occurs in a frame affects the subsequent frames according to the frame type [7], [12], [13]. Distortion severity depends on the amount and the type of the lost data [1].

D. The Loss Models for IP Networks

As presented in our previous works [1], [3], [14], the degradation of a video sequence is evaluated by combining two loss models that are common for IP networks—the identical independent distribution (IID) model and the burst model. The network packet loss simulation is modeled by a two-state Markovian chain, which includes a bursty loss, distances between loss bursts, a single packet loss, and no loss [7], [15]–[17]. In our study, we set the packet loss rate to be 10^{-3}, and derive the probabilities of a correct and an incorrect packet reception from the model.

E. Scene Characteristics

The amount and visibility of distortions in a video scene strongly depend on the actual video content and the human visual system (HVS). To emulate the HVS the following parameters are derived for the video content: a) spatial activity; b) temporal activity; c) spatial-temporal interaction; d) contrast; e) focus of attention and f) colorfulness.

Our study focuses on video transmission over unreliable user datagram protocol over IP (UDP/IP) networks which deploy a network QoS mechanism. UDP is a lightweight transmission protocol that is ideal for real-time video streaming, since it is not loaded with any congestion control or retransmission mechanisms. The playback rate is guaranteed to be constant and smooth at all times, without the constraint of a congestion window structure. Also, UDP-based protocols have a more pronounced advantage over transmission control protocol (TCP) based protocols in competing for network bandwidth. This is because any influx of UDP traffic into the network invokes the congestion control mechanism of existing TCP connections, thereby surrendering the bandwidth to the greedy UDP streams for the sake of maintaining network stability. In a previous work [1] we used several quality measures to quantify the degradation of a video sequence due to losses in an IP communication network. The IID model and the burst model were implemented to simulate the IP network. In addition, in some of our other works [3], [14], subjective tests were used to demonstrate the impact of the coding bit rate, packet loss and video complexity on the perceived video quality. In those papers we emulated the IP loss model as a two state Markovian chain. We analyzed the subjective tests in [3], [14] using multi dimensional scaling (MDS) and a data mining tool. The video content was previously described in [3], [14] using only two parameters, namely spatial and temporal activities, which were measured for the original video sequences (common intermediate format (CIF) resolution, 25 fps, 4 Mbps).

In the present work we parameterize the video content by extracting perceived video content descriptors from all the encoded sequences. These descriptors enable us to generalize the scheme selection problem by describing the video stream. We implement a simplified subjective test on a large group of subjects, and collect solid information for the analysis. We use the statistical F-test (ANOVA) to determine the significance of these results. Finally, based on canonical discriminant analysis (CDA), we implement a statistical classifier to predict the scene quality.

We analyzed viewers’ judgment of video streams with varying spatial and temporal activity levels as well as different levels of contrast and colorfulness. The video streams were processed into two resolutions (CIF and quarter common intermediate format (QCIF)), two frame rates (12.5 and 25 fps), and three bit rates or levels of bit-wise compression (0.5, 1, and 2 Mbps). One type of video sequence having a bit rate of 4 Mbps (the lowest compression) was used as the control group. The perceived video quality was derived from simplified subjective tests on a large group of subjects. We focused on quality optimization, based on the hypothesis that among all the various options to achieve a target bit rate for QoS video transmission and for known video content, there is an encoding scheme that maximizes the user QoE.

The paper continues as follows: Section II addresses related work. In Section III we describe the parameter space and the experimental methods. Section IV introduces the test setup and procedure. Data analysis and results are discussed in Section V. Section VI draws conclusions and suggests related topics for future work.

II. RELATED WORK

Several encoding schemes were suggested in [5] and in [6]. In [5] different segments of encoding processes and standards were combined, and in [6] known image and video encoding techniques were overviewed, both in order to analyze the trade off between bit rate and distortion of video streams analytically, based on rate-distortion theory, using pure mathematical distortion measures. While both these studies made a considerable contribution to video coding, they did not deal with the viewer experience, video characteristics, or transmission impairments caused by packet loss.

Unlike references [5] and [6], the work presented in [7] considered loss effects on the quality of the received video stream. This work dealt with rate control for transmission over wireless channels with burst losses, and it suggested estimating the available bit rate using a-priori knowledge on the channel model and a feedback channel. The performance of the system in [7] was evaluated, similar to all other rate-distortion studies, using pure mathematical measures—PSNR and PLR (this PLR was not a channel parameter, but the ratio of total number of lost packets due to bit errors to total number of packets). However, viewers’ perceived quality was not used for performance evaluation, nor the characteristics of the video streams used for making encoding decisions.

We choose to use a standard available encoder (MPEG2), and transmit with a constant bit rate and a constant compression algorithm throughout the video stream or the scene (since we assume short video streams or scenes with approximately constant characteristics). Our objective is to identify the optimal
encoding scheme that maximizes the viewer-perceived quality. This approach allows the media content to adapt to diverse network characteristics in order to meet the resource constraints such as PLR.

Subjective and objective video quality assessments were addressed in [9], [12] and [18]. Reference [9] used a neural network approach to design an objective quality assessment algorithm for MPEG2 video streams without decoding. The joint impact of loss rate, bit rate and frame rate on the quality of the video stream was studied in [9], and was found to affect the quality, as well as the number of consecutive lost packets (the burst size), and the ratio of the encoded intra macro-blocks to inter macro-blocks, which implies on the number of scene changes or other major changes along the video. Due to the nature of their study, other video characteristic parameters, such as spatial activities and contrast, were not considered and were not found to affect the video quality, whereas our results show they do affect the video quality. In addition, the authors used loss rates of 0, 1, 2, 5 and 10%, which are (except 0%) very high for commercial video streaming, that should be in the range of 0.1% to 0.001% [10].

Reference [12] focused on predicting packet loss visibility in MPEG2 video streams. The authors classified each packet loss as visible or not, and predicted the probability that a packet loss was visible. They classified the measurements factors into two types: content-independent factors and content-specific factors. For content-independent factors, they measured the influence of a single isolated packet loss and the visible artifacts. This study did not aim to address the typical packet loss scenario, which includes a mixture of a single loss model and a burst loss model that are distributed randomly along the bit stream. For content-specific factors, they derived motion descriptors from the motion vectors of MPEG2, but they did not study other parameters that influence the subjective perception, such as spatial activity, colorfulness, brightness and contrast. Another difference between [12] and our research is the subjective test procedure. In the subjective test of [12], the viewers watched a long movie with no reference, and did not grade the video quality, but they reacted to noticeable packet loss artifacts, whereas in our subjective test, the viewers graded the overall quality of short video streams, compared to their original video streams.

In [18], the authors described a general purpose video quality model (VQM) for video systems that span a very wide range of quality and bit rates. This study was adapted as a North American Standard in 2003. This work took into account many objective and subjective parameters that influenced the video streams quality, in order to achieve an objective video quality measure which is highly correlated to viewers’ perception. Their study utilized a reduced-reference model to estimate the overall video quality. In the reduced-referenced model, the assessment of the video quality used low-bandwidth features that were extracted from the source and destination video streams. This is different from our purpose, which is to predict the overall quality of the video stream, by utilizing a non-reference model. In our model, the video quality is estimated prior to video transmission, using the video characteristics, the compression characteristics and the network characteristics.

### III. PARAMETER SPACE AND EXPERIMENTAL METHODS

Subjective video quality rating tests are valuable for designing reliable quality predictors. The standard ITU-R 500 defines three methods for conducting subjective tests [19]:
- DSCQS—Double stimulus continuous quality method
- DSIS—Double stimulus impairment scale method
- SSCQE—Single stimulus continuous quality evaluation

We used a modified version of the DSIS method, in which the viewers watch a pair of video streams (with the same format)—a test scene and a control scene, where the order of the sequences within the pair is not constant and is not known to the viewers. The distortion of the test sequence is graded in a continuous scale between 0 and 100.

Five test sequences were selected from phase one of the video quality experts group (VQEG) [20] on the basis of their different characteristics of temporal and spatial activity, contrast and colorfulness (see Fig. 1). The test sequences were resized into a CIF resolution (resolution and spatial resolution both refer to frame size), which is considered in our tests to be the original frame size. Each original stream was encoded with a different frame-rate, frame size (resolution or spatial resolution) and bit rate, which yielded thirteen different sequences with different encoding schemes for each scene. The parameters that were measured for each of the thirteen generated sequences define the content descriptor of the sequences.

In general, by representing all the input-related parameters (e.g., video content, compression, transmission) with a vector \( X \) of size \( d \) (based on the subjective tests), we can obtain a set of observation pairs of labeled samples \( P_i = \{ X_i, Q_i \} \), where \( Q_i \) is the quality class related to the vector of input parameters \( X_i \). We now demonstrate that by using a pattern classifier, we can classify the input vector of parameters \( X_i \) to the quality class \( Q_i \). The video content descriptor contains the following parameters:

#### A. Spatial Activity

Spatial activity is measured using four oriented derivatives for each video stream, which are calculated with directional Sobel

<table>
<thead>
<tr>
<th>Concert: src3</th>
<th>Toy Train: src10</th>
<th>Race Car: src6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Temporal – Low</td>
<td>Temporal – Low</td>
<td>Temporal – High</td>
</tr>
<tr>
<td>Spatial – Medium</td>
<td>Spatial – High</td>
<td>Spatial – Medium</td>
</tr>
</tbody>
</table>

Rugby: src9 Animation: src4

Temporal – High
Spatial – Low

Temporal – Low
Spatial – Low

Fig. 1. Test sequences selected from phase 1 of VQEG, based on their different characteristics.
and is the number of macroblocks in the frame and frames of each stream (2).

16, and sensitivity depends on the adaptation level to local luminance.

contrast [1], do not apply to complex pictures since human eye methods to calculate contrast, such as the Michelson and Weber (SRC10), Concert (SRC3), Animation (SRC4), Car (SRC6) and Rugby (SRC9).

C. Contrast

Contrast is the relative variation of luminance. Traditional methods to calculate contrast, such as the Michelson and Weber contrast [1], do not apply to complex pictures since human eye sensitivity depends on the adaptation level to local luminance.

Therefore, in our study, we calculate the standard deviation (STD) of the pixel values divided by the average pixel values (Average) for each macroblock (MB) in the spatial domain, as shown in (3). The size of each macroblock is 16 × 16, and they overlap since macroblocks do not divide the frames into “objects” and “backgrounds” as the human viewer sees them. Thus, any overlap captures more transitions between objects and backgrounds, which are then added to the statistics of the frame. The average represents the low frequencies, and the standard deviation represents the band pass frequencies.

The macroblock contrast is calculated for each opponent-color channel separately (B-W, R-G, B-Y), as follows:

\[
Contrast = \frac{STD(MB)}{Average(MB)}
\]

The frame contrast for each color channel is the average over the contrast of all the macroblocks in the frame, the total video sequence contrast for each channel is the average of all the values of the frames contrast, and the total stream contrast is the average over the three color channels. Fig. 4 shows the contrast for the different encoding schemes of each scene.

D. Colorfulness

The colorfulness of a video stream is composed of three components: saturation, hue and brightness. The saturation is the pu...
TABLE I
CONTENT-RELATED PARAMETERS

<table>
<thead>
<tr>
<th>Index</th>
<th>Sequence</th>
<th>Spatial Activity</th>
<th>Temporal activity</th>
<th>Contrast</th>
<th>Saturation</th>
<th>Brightness</th>
<th>BW [Mbps]</th>
<th>Frame rate [EPS]</th>
<th>Resolution</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>src1</td>
<td>6.65</td>
<td>68.88</td>
<td>0.24</td>
<td>0.38</td>
<td>125.96</td>
<td>0.5</td>
<td>12.5</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>src1</td>
<td>6.73</td>
<td>59.39</td>
<td>0.24</td>
<td>0.38</td>
<td>126.59</td>
<td>0.5</td>
<td>25</td>
<td>2</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>59</td>
<td>src9</td>
<td>7.01</td>
<td>111.62</td>
<td>0.23</td>
<td>0.30</td>
<td>108.84</td>
<td>2</td>
<td>25</td>
<td>1</td>
</tr>
<tr>
<td>60</td>
<td>src9</td>
<td>6.30</td>
<td>118.50</td>
<td>0.19</td>
<td>0.30</td>
<td>112.96</td>
<td>4</td>
<td>25</td>
<td>2</td>
</tr>
</tbody>
</table>

Fig. 5. The total saturation of each stream.

Fig. 6. The total brightness of each stream.

Fig. 7. The test system.

TABLE II
CONTENT-RELATED PARAMETERS

<table>
<thead>
<tr>
<th>Index</th>
<th>Sequence</th>
<th>Spatial Activity</th>
<th>Temporal activity</th>
<th>Contrast</th>
<th>Saturation</th>
<th>Brightness</th>
<th>BW [Mbps]</th>
<th>Frame rate [EPS]</th>
<th>Resolution</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>src1</td>
<td>6.65</td>
<td>68.88</td>
<td>0.24</td>
<td>0.38</td>
<td>125.96</td>
<td>0.5</td>
<td>12.5</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>src1</td>
<td>6.73</td>
<td>59.39</td>
<td>0.24</td>
<td>0.38</td>
<td>126.59</td>
<td>0.5</td>
<td>25</td>
<td>2</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>59</td>
<td>src9</td>
<td>7.01</td>
<td>111.62</td>
<td>0.23</td>
<td>0.30</td>
<td>108.84</td>
<td>2</td>
<td>25</td>
<td>1</td>
</tr>
<tr>
<td>60</td>
<td>src9</td>
<td>6.30</td>
<td>118.50</td>
<td>0.19</td>
<td>0.30</td>
<td>112.96</td>
<td>4</td>
<td>25</td>
<td>2</td>
</tr>
</tbody>
</table>

Fig. 8. The total saturation of each stream.

Fig. 9. The total brightness of each stream.

Fig. 10. The test system.

E. Content Characteristics

Each sequence is represented by a vector X containing nine parameters which define the space of different ways to achieve a target bit rate, as shown in Table I. BW is the bandwidth of the stream, frames per second (FPS) are the frame rate units, and QCIF and CIF are the resolution options. During the tests, we varied the spatial activity and temporal activity of the video streams.

In order to derive the frame colorfulness, we calculate the average of the saturation and brightness over the frames. The colorfulness of the total video sequence is the average of the colorfulness results over all the frames. Figs. 5 and 6 present the total saturation and brightness of each stream.

The PLR was set to a constant rate of 10^{-3}, since it is the upper bound of packet loss probability for video streaming [10].

IV. TEST SETUP AND PROCEDURE

A. The Test Communication System

The test system consists of a compression unit, a streamer, a network simulator, and a decoder. The streamer we used is the OPTIBASE Media Gateway (MGW 2000), which is a TV streaming platform for IP networks [23]. The packet loss model was incorporated into the communication network using the RadCom Internet Simulator [24]. This simulator emulates the behavior of wide area network (WAN) links, and introduces impairments such as latency, jitter and packet loss [21].

The test scenes were first transmitted via a PC to the MGW. Then they were encoded into the MPEG2 streams with a CBR, and finally streamed over the IP network. Each encoded stream was transmitted via the RadCom internet simulator, which applies the two-state Markovian packet loss model. The video was received and stored in MPEG2 format in the receiver PC, as depicted in Fig. 7.

As mentioned earlier, to achieve different encoding schemes, each original sequence was encoded with different frame-rates, frame sizes (resolutions) and bit rates as presented in Table II. These original test scenes were compressed into MPEG2 format using a CBR that was set to be the target bit rate. All the encoded streams were reconstructed to the original frame size and rate before starting the subjective tests. To overcome aliasing in the spatial domain and to improve scene visibility, a bicubic interpolation was used to restore the original CIF resolution to those streams which were encoded with QCIF resolution.
TABLE II

METHODS FOR OBTAINING THE TARGET BIT RATE (FRAME RATE, FRAME SIZE AND QUANTIZATION)

<table>
<thead>
<tr>
<th>Scene</th>
<th>Resolution</th>
<th>Frame Size</th>
<th>FPS</th>
<th>Pixel Depth</th>
<th>Original Bit Rate (MB)</th>
<th>Compression Ratio (Q)</th>
<th>Allocated Bandwidth</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>CIF</td>
<td>288</td>
<td>352</td>
<td>25</td>
<td>16</td>
<td>38.67</td>
<td>10:1</td>
</tr>
<tr>
<td>2</td>
<td>CIF</td>
<td>288</td>
<td>352</td>
<td>25</td>
<td>16</td>
<td>38.67</td>
<td>20:1</td>
</tr>
<tr>
<td>3</td>
<td>CIF</td>
<td>288</td>
<td>352</td>
<td>25</td>
<td>16</td>
<td>38.67</td>
<td>40:1</td>
</tr>
<tr>
<td>4</td>
<td>CIF</td>
<td>288</td>
<td>352</td>
<td>25</td>
<td>16</td>
<td>38.67</td>
<td>80:1</td>
</tr>
<tr>
<td>5</td>
<td>CIF</td>
<td>288</td>
<td>352</td>
<td>12.5</td>
<td>16</td>
<td>19.34</td>
<td>10:1</td>
</tr>
<tr>
<td>6</td>
<td>CIF</td>
<td>288</td>
<td>352</td>
<td>12.5</td>
<td>16</td>
<td>19.34</td>
<td>20:1</td>
</tr>
<tr>
<td>7</td>
<td>CIF</td>
<td>288</td>
<td>352</td>
<td>12.5</td>
<td>16</td>
<td>19.34</td>
<td>40:1</td>
</tr>
<tr>
<td>8</td>
<td>QCIF</td>
<td>144</td>
<td>176</td>
<td>25</td>
<td>16</td>
<td>9.67</td>
<td>5:1</td>
</tr>
<tr>
<td>9</td>
<td>QCIF</td>
<td>144</td>
<td>176</td>
<td>25</td>
<td>16</td>
<td>9.67</td>
<td>10:1</td>
</tr>
<tr>
<td>10</td>
<td>QCIF</td>
<td>144</td>
<td>176</td>
<td>25</td>
<td>16</td>
<td>9.67</td>
<td>20:1</td>
</tr>
<tr>
<td>11</td>
<td>QCIF</td>
<td>144</td>
<td>176</td>
<td>12.5</td>
<td>16</td>
<td>4.83</td>
<td>2.5:1</td>
</tr>
<tr>
<td>12</td>
<td>QCIF</td>
<td>144</td>
<td>176</td>
<td>12.5</td>
<td>16</td>
<td>4.83</td>
<td>5:1</td>
</tr>
<tr>
<td>13</td>
<td>QCIF</td>
<td>144</td>
<td>176</td>
<td>12.5</td>
<td>16</td>
<td>4.83</td>
<td>10:1</td>
</tr>
</tbody>
</table>

In the temporal domain, the reconstruction from 12.5 fps to 25 fps was made by using a simplified bilinear interpolation. The reference video stream was encoded with CIF resolution, a frame-rate of 25 fps and a 4 Mbps bit rate, in order to achieve the highest quality after compression, and it was not transmitted through the lossy network simulation, in order to get the best achievable reference video stream for all the other video streams. The viewers watched all thirteen test sequences and compared them to the reference sequence.

B. The Subjective Tests

For these tests, subjects were shown both the distorted test sequence and the reference sequence. The order of the sequences within the pairs was not constant and was unknown to the viewers. The subjects were first asked to decide which scene is better, and then to grade the similarity percentage of the worst scene compared to the better one. This approach was chosen in order to improve the classification of the quality rank of each distorted scene. A visual basic script guided the subjects via the test scenarios, and the subjects’ preferences were saved in a database for analysis. The test scenarios are described in our previous work [3], [14].

C. Subjective Data Collection and Preliminary Results

Results were collected from 93 subjects. The subjects were software engineering students who were well acquainted with MPEG behavior. One record was stored for each scene. Each record contained information on the subject, the movie, the scene under test, the selected scene and the percentage distance of the unselected scene from the selected scene. The data files from the 93 subjects contained 6045 records which described the quality space of the test sequences. Fig. 8 and Table III present the average grades of the subjective tests for each sequence of each scene.

It can be seen from Table III that the scenes SRC10 (Toy-Train) and SRC3 (Concert) scored the highest grade in the second encoding scheme (CIF, 12.5 fps, 1 Mbps). They both had low temporal activity, SRC10 with high spatial activity, and SRC3 with medium spatial activity. SRC4 (Animation) scored the highest grade in the fourth encoding scheme (CIF, 25 fps, 0.5 Mbps) which reduced the bit rate only by MPEG compression. This scene has low temporal and spatial activities. Scenes SRC6 (Race-Car) and SRC9 (Rugby) are both sports-oriented, thus they have high temporal activity. SRC9 has low spatial activity, and SRC6 has medium spatial activity. The highest grade for these two scenes was achieved in the twelfth encoding scheme (QCIF, 25 fps, 1 Mbps).

Based on these results, we can conclude that for high spatial activity the CIF resolution is preferred, whereas for high temporal activity, the full frame-rate is preferred. Only for SRC4 did the preferred encoding scheme have full resolution and 25 fps. As expected, none of the preferred encoding schemes had a bit rate of 2 or 4 Mbps, since many errors resulted from packet loss (as the bit rate became higher, more meaningful packets were lost for the same PLR [12]). Thus, a video stream with high bandwidth would have a rather low quality for constant PLR with respect to lower bit rates, but not necessarily the worst quality. On the other hand, the disadvantage of reducing the frame rate is that the distances between the frames that are not discarded is bigger, thus the motion vectors are longer, the differences between the blocks grow, and each loss causes more...
TABLE III
AVERAGE GRADES OF THE SUBJECTIVE TESTS

<table>
<thead>
<tr>
<th>Encoding scheme</th>
<th>Scene Sequence</th>
<th>SRC10</th>
<th>SRC9</th>
<th>SRC6</th>
<th>SRC4</th>
<th>SRC3</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5Mbps</td>
<td>12.5fps CIF</td>
<td>60</td>
<td>44</td>
<td>58</td>
<td>66</td>
<td>59</td>
</tr>
<tr>
<td>1Mbps</td>
<td>12.5fps CIF</td>
<td>71</td>
<td>50</td>
<td>64</td>
<td>64</td>
<td>82</td>
</tr>
<tr>
<td>2Mbps</td>
<td>12.5fps CIF</td>
<td>66</td>
<td>59</td>
<td>69</td>
<td>54</td>
<td>57</td>
</tr>
<tr>
<td>0.5Mbps</td>
<td>25fps CIF</td>
<td>4</td>
<td>59</td>
<td>50</td>
<td>45</td>
<td>71</td>
</tr>
<tr>
<td>1Mbps</td>
<td>25fps CIF</td>
<td>5</td>
<td>67</td>
<td>60</td>
<td>64</td>
<td>64</td>
</tr>
<tr>
<td>2Mbps</td>
<td>25fps CIF</td>
<td>5</td>
<td>67</td>
<td>60</td>
<td>64</td>
<td>64</td>
</tr>
<tr>
<td>4Mbps</td>
<td>25fps CIF</td>
<td>7</td>
<td>64</td>
<td>62</td>
<td>60</td>
<td>60</td>
</tr>
<tr>
<td>0.5Mbps</td>
<td>12.5fps QCIF</td>
<td>8</td>
<td>59</td>
<td>48</td>
<td>65</td>
<td>54</td>
</tr>
<tr>
<td>1Mbps</td>
<td>12.5fps QCIF</td>
<td>9</td>
<td>63</td>
<td>59</td>
<td>74</td>
<td>58</td>
</tr>
<tr>
<td>2Mbps</td>
<td>12.5fps QCIF</td>
<td>10</td>
<td>58</td>
<td>68</td>
<td>61</td>
<td>57</td>
</tr>
<tr>
<td>0.5Mbps</td>
<td>25fps QCIF</td>
<td>11</td>
<td>62</td>
<td>62</td>
<td>66</td>
<td>53</td>
</tr>
<tr>
<td>1Mbps</td>
<td>25fps QCIF</td>
<td>12</td>
<td>62</td>
<td>72</td>
<td>78</td>
<td>62</td>
</tr>
<tr>
<td>2Mbps</td>
<td>25fps QCIF</td>
<td>13</td>
<td>63</td>
<td>63</td>
<td>51</td>
<td>56</td>
</tr>
</tbody>
</table>

TABLE IV
SCENES CHARACTERISTICS AND THEIR PREFERRED ENCODING SCHEME

<table>
<thead>
<tr>
<th>Scene</th>
<th>Characteristic</th>
<th>Spatial activity</th>
<th>Temporal activity</th>
<th>Contrast</th>
<th>Colorfulness: saturation</th>
<th>Colorfulness: brightness</th>
<th>Preferred encoding</th>
</tr>
</thead>
<tbody>
<tr>
<td>Toy-Train (SRC10)</td>
<td>High</td>
<td>Low</td>
<td>High</td>
<td>High</td>
<td>High</td>
<td>Medium</td>
<td>(2) CIF, 12.5 fps, 1 Mbps</td>
</tr>
<tr>
<td>Concert (SRC3)</td>
<td>Medium</td>
<td>Low</td>
<td>Low</td>
<td>High</td>
<td>High</td>
<td>Medium</td>
<td>(2) CIF, 12.5 fps, 1 Mbps</td>
</tr>
<tr>
<td>Animation (SRC4)</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>High</td>
<td>Low</td>
<td>Low</td>
<td>(4) CIF, 25 fps, 0.5 Mbps</td>
</tr>
<tr>
<td>Race-Car (SRC6)</td>
<td>Medium</td>
<td>High</td>
<td>Medium</td>
<td>Low</td>
<td>High</td>
<td>Medium</td>
<td>(12) QCIF, 25fps, 1 Mbps</td>
</tr>
<tr>
<td>Rugby (SRC9)</td>
<td>Low</td>
<td>High</td>
<td>Medium</td>
<td>Medium</td>
<td>Medium</td>
<td>Medium</td>
<td>(12) QCIF, 25fps, 1 Mbps</td>
</tr>
</tbody>
</table>

V. DATA ANALYSIS

A. Principle Component Analysis (PCA)

In order to classify the subjective test results into classes of quality, we assume a $d$ dimension normal probability density of the results, which is formulated as:

$$p(x) = \frac{1}{2\pi^{d/2}|\Sigma|^{1/2}} e^{-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)}$$

(4)

where the mean $\mu$ is a nine-dimensional vector, $\Sigma$ is a $9 \times 9$ covariance matrix, and $|\Sigma|$ is the determinant of $\Sigma$. The density function $p(x)$ is governed by the parameters $\mu$ and $\Sigma$ which satisfy:

$$\mu = \mathbb{E}(x)$$

and

$$\sum = \mathbb{E}[(x-\mu)(x-\mu)^T]$$

(5)

where $\mathbb{E}$ is the expectance, and the exponent is a Mahalanobis distance $\Delta$ from $x$ to $\mu$:

$$\Delta^2 = (x-\mu)^T\Sigma^{-1}(x-\mu)$$

(6)

In the bar histogram in Fig. 9, we can see that the first three principle components explain roughly more than two-thirds of the total variability in the standardized ratings. The continuous graph is the sum of the histogram bars [25].

B. Analysis of Variance (ANOVA)

Based on PCA and by examining the preferred encoding scheme for each scene with regard to the scene characteristics (Table IV), we found that temporal activity and spatial activity damage. Moreover, reconstructing the frame-rate and the resolution to the original values resulted in error propagation to the reconstructed frames.

Our results coincide in general with those of [4], where lower spatial resolution and lower frame rate were preferred over high quantization, where quantization is the MPEG compression control parameter. Nevertheless, there was no packet loss at all in [4], thus as the bit rate increased, the perceived quality increased as well. This is not the case in our study, where higher bit rate results in more packet loss, which degrades the overall perceived image quality due to data loss and the error propagation nature of the compressed video sequences.
are the two major characteristics that affect video stream quality. To determine the significance of these results, we analyzed the variances between the groups, which would identify those factors which cause differences in the means. ANOVA can identify any significant factor that causes difference in variance between groups [26]. A repeated measures ANOVA [27] entails using the same subjects for different parts of the test. The advantage of this method is that fewer subjects are needed, since the same subjects’ preferences are tested on the different parts of the test parameters.

The ANOVA output is $F$, which is the probability to reject the null hypothesis when the alternative hypothesis is correct. $F$ is equals to the ratio between the real variation of the group averages, and the variation of the estimated averages of the results (8). If the null hypothesis is not rejected, $F$ has a value close to 1. A big value of $F$ (much higher than 1) would mean that the tested factors have influenced the results, and that the rejection of the null hypothesis is correct [26].

$$F = \frac{\text{Found variation of the group averages}}{\text{Expected variation of the group averages}}$$ (7)

The probability to decline the null hypothesis by mistake is $p$, meaning that the null hypothesis is true, but that the results indicate otherwise. A very low value of $p$ (lower than $10^{-3}$) means that the null hypothesis is probably incorrect [26].

Five scenes were graded, each with thirteen encoding schemes (Table III). Each of 93 viewers graded each encoded and transmitted stream on a scale ranging from 0 to 100. We hypothesized that the different characteristics of the scenes would affect the viewers’ perceived quality, and that there would be a relationship between the scene characteristics and the preferred encoding scheme. For example, if a scene has high spatial activity, the viewers would prefer an encoding scheme with high resolution, and if a scene has high temporal activity, the viewers would prefer an encoding scheme with low resolution. The scenes were categorized into groups for each of the video characteristics presented in Table V. For each encoding scheme in each group, we averaged the scores for each viewer.

1) Spatial Activity versus Resolution: A two-way ANOVA (spatial activity(2) × resolution(2)) computed on the rating score of the subjective observers across all bit rates yielded a significant interaction between spatial activity and resolution $\left[F_{(1,81)} = 5.23; p = 0.025\right]$, with subjects evidencing a preference for the CIF resolution under high spatial activity. Post hoc analysis showed significant differences between the two resolution values under conditions of high spatial activity, but did not evidence significant differences for lower spatial activity (see Fig. 10). These results are consistent with HVS behavior, since the high resolution scene captured the high spatial frequencies, which are important in a high spatial-activity scene. When the scene contained low spatial activity, the subjects did not have significant preferences, and a high resolution scene could be replaced with a low resolution scene in order to save resources. In general, subjects gave higher scores to scenes that contained high spatial activity for both CIF and QCIF resolution, indicating that scenes with high spatial activity are preferred by the HVS. We also tested the interplay between spatial activity and resolution separately for different bit rates. A two-way ANOVA (spatial activity(2) × resolution(2)) computed on the rating score of the subjective observers for a bit rate of 1 Mbps $\left[F_{(1,82)} = 19.4; p < 0.001\right]$. Post hoc analysis yielded a noticeable difference between the resolution values, where at high spatial activity observers preferred the CIF resolution, and at low spatial activity they preferred the QCIF resolution (see Fig. 11). This result is reasonable, since the compression error for compressing at the same bit rate is lower for QCIF resolution with low spatial activity than for CIF resolution, so that a resolution reduction in low spatial activity actually improves the subjective quality rating. The ANOVA results $F$ and $p$ only determine if the null hypothesis is incorrect, but not if the alternative hypothesis is correct. The subjective test

<table>
<thead>
<tr>
<th>Scene</th>
<th>Spatial Activity</th>
<th>Temporal Activity</th>
<th>Contrast</th>
<th>Saturation</th>
<th>Brightness</th>
</tr>
</thead>
<tbody>
<tr>
<td>SRC3</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>SRC4</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>SRC6</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>SRC9</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>SRC10</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Average over SRCs:</td>
<td>10 3.6 4.9</td>
<td>6.9 10 3.4</td>
<td>10 3.6 9</td>
<td>3.4 10</td>
<td>9 6 10</td>
</tr>
</tbody>
</table>

Fig. 10. Resolution versus activity, over all the bit rates.
results should be further analyzed, in order to determine the combined effect of encoding scheme parameters and scene characteristics parameters on the viewers’ preferences, by using post hoc analysis. This analysis is performed only when the ANOVA value $F$ is high, meaning there is a significant difference in the means of the parameters. It reveals not only if a scene characteristics parameter is significant, but also how it affects the subjects’ preferences.

For a target bit rate of 1 Mbps with CIF resolution, the subjects’ preferences coincided with scenes with high and low spatial activity, due to the deep MPEG2 compression which removes the high frequencies. For the low resolution scene (QCIF), the subjects gave higher scores for the scenes with low spatial activity, which means that the distortions from the MPEG2 compression and error propagation due to transmission errors are greater in scenes with high spatial activity. On the other hand, for a bit rate of 2 Mbps (Fig. 12), at high spatial activity, observers preferred the QCIF resolution and at low spatial activity no significant preference was found $[F(1,82) = 32.31; p < 0.0001]$. These results can be explained by the fact that for scenes with high spatial activity, the CIF resolution suffers from deeper MPEG2 compression than the QCIF resolution, since the latter contains fewer pixels for the same target bit rate.

2) Temporal Activity versus Frame Rate: We compared the subjects’ responses to scenes with low temporal activity to those with high temporal activity, to see how the frame rate could affect the subjects’ perception of video scenes with low temporal activity. A two-way ANOVA (temporal activity(2) × frame rate(2)) computed on average scores over all resolutions and bit rates yielded a statistically significant main factor for frame rate $[F(1,81) = 7.87; p = 0.006]$. No significant main factor was found for temporal activity, and the interaction of factors was not significant. Post hoc analysis yielded that the frame rate significantly affected evaluation scores at low temporal activity, but did not significantly affect scores at high temporal activity. At the low frame rate, there was no significant difference in scores based on temporal activity, whereas at the high frame rate there was (see Fig. 13). These finding seems reasonable, since lowering the frame rate of low temporal activity scenes should not significantly degrade the perceived quality very much. Furthermore, MPEG2 compression uses a lower quantization to reach the target bit rate, in which case a decrease in the frame rate would improve the subjects’ preferences for the scenes with low temporal activity.

To evaluate the impact of the frame rate on the subjective results, we considered the bit rates separately. At a bit rate of 0.5 Mbps, hence, the interaction between temporal activity and frame rate was not found to be significant (Fig. 14(a)). Both temporal activity and frame rate were, individually significant main factors in the observer scores $[F(1,81) = 20.41; p < 0.001]$ and $F(1,81) = 14.87; p < 0.001$, respectively. Furthermore, high compression yielded low quality scenes that masked the degradation due to the frame rate, and thus, the subjects’ preferences were similar.

When we included only the video sequences with the lowest compression, namely a bit rate of 2 Mbps, in the analysis, no significant main factors were found, although the interaction between temporal activity and frame rate was statistically significant $[F(1,81) = 64.7; p < 0.0001]$. Post hoc analysis showed that all combinations were statistically significant. For scenes
with high temporal activity, the lower frame rate was found by the subjects to be more effective and for scenes with low temporal activity, the higher frame rate was preferred, both of which are shown in Fig. 14(b).

This last result may seem surprising, as we would expect that scenes with high temporal activity would get a higher score with a higher frame rate, owing to the fact that lowering the frame rate degrades the smoothness of the motion in the video stream. However, since the high frame-rate scenes are compressed more than the low frame-rate scenes in order to achieve the same target bit rate (2 Mbps), the compression causes more dependencies between elements such as reference blocks and motion vectors, and consequently each loss causes more damage to the video quality. Moreover, the lost packets might include more meaningful data than other packets.

3) Contrast versus Compression Rate: To test the hypothesis that at high compression rates, high contrast negatively affects the subjects’ evaluation of the video sequences, we computed a two-way ANOVA (contrast(2) × bit rate(3)) on the average score (see Fig. 15). Our initial results showed a significant interaction between the variables ($F_{2,144} = 30.36; p < 0.0001$). However, post hoc analysis showed that there was a significant difference in the scores between the lower bit rates (0.5 Mbps) and the higher bit rates (1 and 2 Mbps), but not between bit rates of 1 and 2 Mbps at both high and low contrast. For the low bit rate (0.5 Mbps), the contrast significantly affected the score, while low contrast scenes received higher scores than high contrast scenes, because as expected, the degradation due to MPEG2 compression was higher and more noticeable for high contrast. On the other hand, at bit rates of 1 and 2 Mbps, the contrast was not found to significantly affect observer scores, because the effects of compression and packet loss were similar between the scenes.

4) Summary of the ANOVA Test Results: The hypothesis that the subjects’ preferences for scene quality depend on the scene characteristics, on the compression, and on the channel behavior was supported by the results of the statistical $F$-test (ANOVA). Based on these results we can conclude the following:

1) Resolution is a significant factor for scenes with high spatial activity, but not for scenes with low spatial activity.
2) Temporal activity is a significant factor for the frame rate, and subjects’ preferences also differ according to the compression rate.
3) High contrast negatively affects subjects’ evaluations of scenes at high compression rates.

C. Canonical Discriminant Analysis

Since there is a dependency between all the scene characteristics and the preferred encoding scheme, an overall dependency analysis was used to exploit the effect of the scene characteristics on the subjects’ perceived quality of the video streams. As explained in Section III, by representing all the input-related parameters with a vector $X$ containing $d$ parameters, we obtain a set of observation pairs of labeled samples $P = \{X_i, Q_i\}$, where $Q_i$ is the quality class related to the vector of input parameters $X_i$. The vector $X$ contains the video content descriptor, the compression descriptor and the network descriptor. We now consider quality classification at three and four quality levels. In the three-levels classifier, we classify an input vector of parameters $X$ to the following three quality levels: $Q_1$ representing low quality, $Q_2$ representing mid quality, and $Q_3$ representing high quality. For the four-level classification, $Q_1$ represents low quality, $Q_2$—mid low quality,
TABLE VI
CHI-SQUARE WITH SUCCESSIVE ROOTS REMOVED

<table>
<thead>
<tr>
<th>Removed</th>
<th>Eigenvalue</th>
<th>Chi-sqr</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>136.5676</td>
<td>0.0000</td>
</tr>
<tr>
<td>1</td>
<td>0.3261</td>
<td>16.7941</td>
<td>0.0188</td>
</tr>
<tr>
<td>2</td>
<td>0.0943</td>
<td>5.4063</td>
<td>0.4928</td>
</tr>
</tbody>
</table>

$Q_3$—mid high quality and $Q_4$—high quality. These quality classes were derived from the subjective tests.

To classify quality, we used linear discriminant analysis, which implements a dimension-reduction technique. This technique is related to principal component analysis and canonical correlation (also called canonical discriminant analysis (CDA)). CDA derives the canonical coefficients as a one-way multiple analysis of variance (MANOVA) and finds linear combinations of the quantitative variables. It also provides maximal separation between the classes or groups, in a manner similar to summarizing the total variances of the principal components. The outputs are the canonical coefficients and the scored canonical variables. Bartlatt’s approximate Chi-squared statistic is used for testing the canonical correlation coefficients [28]. The decisions on quality membership in our CDA classifier are based on $d$ dimension normal probability density as in (5).

1) Classification Results: Based on the subjective tests, we obtained 65 observation pairs of labeled samples $P = \{X_1, Q_1\}$ for the three and four quality levels. The CDA classifier has a significance level of $\alpha = 5\%$. The test of significance gives a $p$-value lower than the $\alpha$-level as shown in Tables VI(a) and (b), were $\alpha$ represents the significance level of a test which is the probability that the null hypothesis will be falsely rejected when it is true.

Our classifiers are shown in the scatter plots in Figs. 16(a) and (b). It can be seen that in the three-level classifier that false classifications could accrue from neighbor quality classes, from $Q_1$ to $Q_2$ and from $Q_3$ to $Q_2$, but not from $Q_1$ to $Q_3$ or vice versa. For the four-level classifier it can be seen that class $Q_4$ is spread over the quality classes $Q_2$ and $Q_3$, which could cause misclassifications between distant classes.

To validate our classifier, we used a semi-cross validation technique. For the design stage we used 50 labeled samples from which 15 samples were used for validation. We repeated this process several times by selecting different samples randomly for each stage. The three-level classifier achieved better classification results than the four-level classifier. The three-level classifier achieved average hit-rates of 86% (13 out of 15 samples), and the four-level classifier achieved average hit-rates of 80% (12 out of 15 samples). By analyzing the nature of the missed samples we found that classifications errors occurred only from neighbor quality classes in the two classifiers.

D. The Prediction Model

To predict the perceived video quality prior to video transmission we introduced a new hypothetic reference circuit (HRC), as shown in Fig. 17. The HRC contains a measurement subsystem that measures and derives a set of parameters which are related to the perceived video quality. This set of parameters includes the content descriptor, the compression descriptor and...
the network descriptor. Based on the set of parameters and on the CDA statistical classifier, the objective quality rating subsystem was used to evaluate subjective video quality. To identify the optimal encoding scheme that maximizes viewer-perceived quality, the process was repeated with different encoding schemes. This method allows the media content to adapt to diverse network characteristics so as to meet the resource constraints such as bandwidth and PLR.

VI. SUMMARY AND FUTURE WORK

The experimental results presented in this paper indicate that the video encoding scheme, the channel behavior and the video content all influence the quality of a video stream as experienced by viewers. We generalized the classification problem by describing the video stream and the channel quantitatively such that prior knowledge of the video content was not required. Based on intensive subjective testing, we obtained a set of observation pairs of labeled samples that mapped the parametric space into quality ranks. We demonstrated that when jointly considering the impact of the coding bit rate, the packet loss ratio, and the video characteristics, we can derive an optimal encoding scheme that maximizes viewer-perceived quality. The statistical F-test (ANOVA) was employed to determine the significance of these results.

From a practical point of view, we showed that it is feasible to classify and predict the subjective quality of video stream from a set of parameters by using a linear discriminant analysis technique known as CDA. A new HRC was presented that determines the optimal encoding scheme and the required content adaptation to diverse network characteristics. Integrating this knowledge into existing transmission networks can assist video application service providers in choosing the transmission methods that best suit the network conditions and the video sequence characteristics.

A future stage of this research would be to experiment on a wider variety of scenes, in order to achieve a wider range of values of scene characteristics. For this purpose, each encoded stream would have to be transmitted through the simulation network several times, which would enable a broader distribution of the locations of packet loss [12]. We also plan to expand the subjective tests using PLR values of 10^{-4} and 10^{-8}, which are found to be very common in today’s networks. Another further research step is related to QoE scale selection in the emerging H-264 Scalable Video Coding standard [29]. From our results it would seem that selecting the video scale according to the channel characteristics and the video content could improve the viewers’ quality of experience.

REFERENCES
Ron Shmueli received the Ph.D. and B.Sc. degrees in electrical and computer engineering from Ben-Gurion University and M.Sc. degree from Tel-Aviv University. From 1992–1999 he led the Intelligent Network department in Tadiran Telecommunication Ltd and from 2000–2002 he was the CEO of HOMEK Network Solutions Inc. Currently, he is a faculty member at the Electrical Engineering and Computer Department at AFEKA-Academic College of Engineering in Tel-Aviv, Israel.

Ofer Hadar (M’97) received the BSc, the MSc (cum laude) and the PhD degrees from the Ben-Gurion University of the Negev, Israel, in 1990, 1992, and 1997, respectively, all in electrical and computer engineering. The prestigious Clore Fellowship supported his PhD studies. His PhD dissertation dealt with the effects of vibrations and motion on image quality and target acquisition. From August 1996 to February 1997, he was with CREOL at Central Florida University, Orlando, FL, as a research visiting scientist, working on angular dependence of sampling MTF and over-sampling MTF. From October 1997 to March 1999, he was a post-doctoral fellow in the Department of Computer Science at the Technion-Israel Institute of Technology, Haifa. Currently, he is a faculty member at the Communication Systems Engineering Department at Ben-Gurion University of the Negev. His research interests include image compression, video compression, routing in ATM networks, flow control in ATM networks, packet video, transmission of video over IP networks, and video rate smoothing and multiplexing. Dr. Hadar also works as a consultant, and is a member of IEEE and SPIE.

Revital Huber received her BS degree in electrical engineering from the Technion- Israel Institute of Technology, Israel, in 1999, and the MSc degree in electrical engineering from the Electrical Engineering Dept. at Ben Gurion University of the Negev, Israel, in 2004. She has served as an academic professional officer in the Israel Air Force, and was a teaching assistant at Ben Gurion University of the Negev. She is currently working toward her PhD degree in Communication Systems Engineering Dept. at Ben Gurion University of the Negev.

Masha Maltz is a lecturer at Ben-Gurion University of the Negev in Beer Sheva, Israel. She received her BA in economics and her MBA in computer methodology from the City University of New York in 1978 and 1982, respectively, and her PhD degree in industrial engineering and management from Ben-Gurion University of the Negev in 2001. Her current research interests include target cueing and detection, multispectral imaging, eye movements, and human-computer interaction.

Merav Huber received her BSc degree in electrical engineering from the Technion- Israel Institute of Technology, Israel, in 1999, and the MSc degree in electrical engineering from the Electrical Engineering Dept. at Ben Gurion University of the Negev, Israel, in 2004. She has served as an academic professional officer in the Israel Air Force, and was a teaching assistant at Ben Gurion University of the Negev. She is currently working toward her PhD degree in Communication Systems Engineering Dept. at Ben Gurion University of the Negev.