J. King Saud Univ., Vol. 8, *Eng. Sci.* (2), pp. 267-282 (A.H. 1416/1996)

Ratings of Visual Quality in Digitized Imaging Systems

Awad Kh. Al-Asmari* and Syed Abid A. Naqvi**

**Department of Electrical Engineering and **Department of Mechanical Engineering, College of Engineering, King Saud University, P.O. Box 800, Riyadh 11421, Saudi Arabia*

(Received 13/12/1994; accepted for publication 14/6/1995)

Abstract. This paper describes a study conducted on 40 human subjects to rate six different digitized visual images (3 colored and 3 black & white) on a qualitative scale. The results were compared analysis of variance and pairwise comparison tests. It was found that the subjects responded most favorably to the highest resolution colored image, hence validating the hypothesis that high definition digitized images are superior to the conventional ones. Moreover, it was also found that 5 out of 15 pairs of images were not significantly different from one another, indicating that the images with resolution very close to one another were not being differentiated by the subjects. The findings of this study are of vital significance because it reinforces the objective claims that the high definition images used in high definition television are superior to the low definition images in conventional television.

Introduction

High definition television (HDTV) research is a subject which has gained lot of momentum in the recent years. Several researchers in the area of digitized visual images have addressed different issues within the subject. Some of these studies are described in the following sections.

Special reports [1;2] on the subject update the status of development and use of HDTV in the world. Also, current uses of the system were identified, including the exhibition of still pictures in art galleries. The articles also describe the various difficulties being experienced in the development of a world standard for HDTV. Another report [3] praises the use of HDTV projetion - CRT systems by proving them to be cheaper, less bulky and lighter in weight as compared to the large direct - view CRT's.

Some researchers [4-8] have shown concerns and addressed different issues pertaining to the international standards on HDTV and universal interchanging of all kinds of images and image sequences [9]. Some of the more advanced aspects such as broadcasting techniques have been addressed. These include: digital HDTV compression techniques [10], digital HDTV terrestrial broadcasting systems [11], tests of the terrestrial broadcasting system [12], methods for accommodation of HDTV terrestrial broadcasting [13], HDTV studio standards [14].

On the human factors side a psychophysically justified bit allocation algorithm for use with subband image coding system was presented [10]. It was found that this algorithm was superior to the minimum mean square algorithm at low bit rates.

Human factors in imaging display designs

Researchers have addressed different aspects of imaging displays, image size [2;15], image quality, target acquisition performance as a function of operator, target and electronic imaging system characteristics [16], image legibility, visual search [17], perception and movement [18;19], and color [1;20;21].

However, visually the main objective of the observer is to rcognize distinctive target patterns against different backgrounds. Some of the factors that help the observer include increased target size, reduced number of irrelevant targets, high contrast, color, large display area, magnification, viewing time etc. Factors relating to the visual system of a person include: their state of adaptation, their spatial acuity and their line of sight. All these factors interact to produce a response of an individual affecting both the speed and accuracy of human performance. An understanding of these factors and their consideration in the design can help predict the system performance more effectively. A recent article [22], identified the human visual system as the key to creating or developing outstanding displays. The critical factors from the visual standpoint relating to display design as described by Rogowitz include the need to recognize the capabilities and limitations of the visual system, luminance perception, spatial vision, flicker and color recognition. The purpose of this study is to decompose an image and then to test the visual quality of each image using human subjects.

Image Decomposition and Reconstruction Procedure

The original image is decomposed into four bands using the sub-band coding (SBC) technique as shown in Fig. 1(a). This technique consists of using two different filters, H, a low-pass filter, and G, a high-pass filter, to decompose the image.

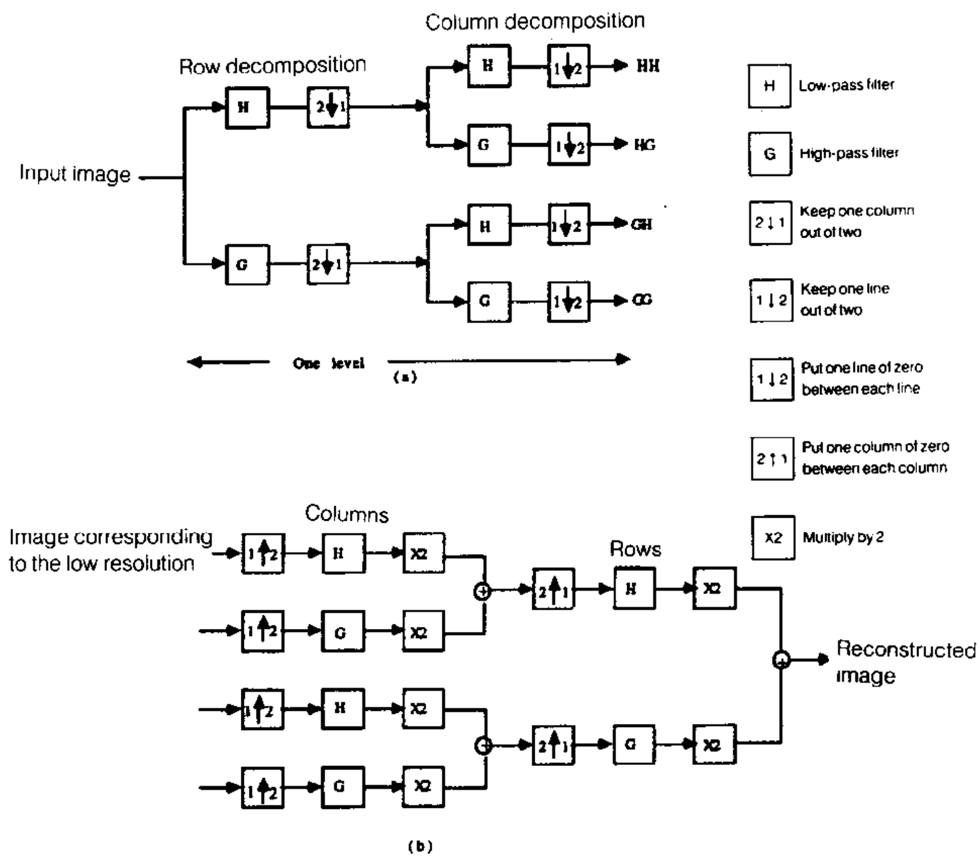

Fig. 1. Sub-band coding of image.
(a) Image decomposition to four bands.
(b) Image reconstruction of the four bands.

The relation between the analysis filters (H and G) and the synthesis filters ($\tilde{H}$ and $\tilde{G}$) for the system shown in Fig. 1 could be given by the following formulas:

$$G(Z) = H(-Z)$$
$$\tilde{H}(Z) = 2H(Z)$$
$$\tilde{G}(Z) = -2G(-Z)$$

where H(Z) is a symmetrical finite inpulse response (FIR) filter of even order. The two filters, H a low-pass filter, and G, a high-pass filter are used to reconstruct the decomposed image. For a more detailed analysis on the quadrature mirror filtering (QMF) and the decomposition of images see [23;24].

For the original image, we used one level of decomposition. The G and H filters are applied to the image in both horizontal and vertical directions as shown in Fig. 1(a). Then, the filter outputs are subsampled by a factor of two, generating three selectively oriented high-pass sub-bands, GG, GH, HG, and a low-pass sub-band HH. The four-band SBC decomposition is shown in Fig. 2.

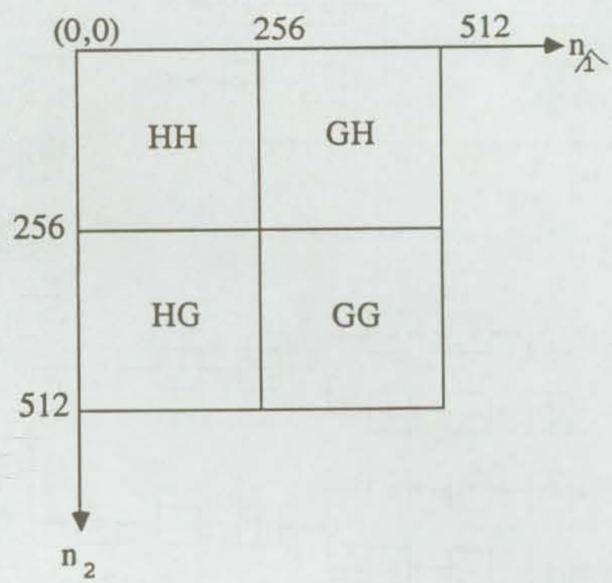

Fig. 2. Sub-band coding image decomposition listed for the four bands in the time domain.

Fig. 3. Original image.

Fig. 4. Sub-band decomposition of original image. The four sub-bands (HH, HG, GH and GG).

The image used in this paper is a digitized color NTSC image with dimensions of 512 × 512 pixels and 8 bits per pixel (bpp). Figure 3 shows the original image decomposed in this manner. The filter used for decomposition is from the QMF family, given in [25], known as (QMF 16A with 16-taps). Figure 4 shows the four bands listed as given in Fig. 2 for the same image. These bands contain 1/4 of the original samples (*i.e.* 256 × 256 pixels each).

QMF sub-band coding algorithms for images have been explored by Vetterli [26], and for use in compression, by Woods and O'Neil [27] and Gharavi and Tabatabai [17]. We have used the QMF and the generalized QMF for sub-band coding of HDTV images [28;29]. In this paper, the main objective is to study the human visual quality and compare it with the quantitative measure. Therefore, we did not design the coding scheme required for compression. We only reconstructed the output image from the base-band HH, which had the most information about the original image. Then, we added all or some of the high-bands GH, HG and GG into HH band. The overall quality of these images was then studied. The reconstruction scheme of the image is presented in Fig. 1(b).

Subjective Evaluation Procedure

It is common to use some form of subjective evaluation to ascertain the quality of the reconstructed images. A number of methods exist, including, binary decision methods and mean opinion scores (MOS). The MOS method as used in this research consisted of 40 university students (21-23 years) rating the quality of images. Each observer was asked to rate the image on a scale of 1 to 5, and the average of the scores was used as the MOS rating of the image. The five point scale, the associated quality and impairment scales, and the evaluation form used by the observers, are shown in Table 1.

Table 1. Subjective evaluation rating scales and evaluation form

Evaluator's Name: ______________________________ Date: __________

Subjective evaluation rating scales

Numerical	Qualitative	Impairment description
5	Excellent	Imperceptible
4	Good	Perceptible but not annoying
3	Fair	Perceptible and slightly annoying
2	Poor	Annoying
1	Bad	Very annoying

Procedure: As each image is displayed, circle the number which best describes the quality of the image.

Image 1	1	2	3	4	5
Image 2	1	2	3	4	5
Image 3	1	2	3	4	5
Image 4	1	2	3	4	5
Image 5	1	2	3	4	5
Image 6	1	2	3	4	5

Six images (3 colored and 3 black and white) were selected to represent several classes of images. Three images (either colored or black and white) were presented separately. The observers were seated at a distance of six times the picture height (to maintain a constant visual angle) and were asked to rate the quality of each image on the scale of 1 to 5 as it was presented on a screen through a tachistoscope, with equal interval and exposure times for all slides. The sequence of the presentations was random except that no image was repeated in a given sequence. Sequences of colored pictures were presented separately from the black and white sequences to avoid bias. The interval between the colored picture testing and black and white testing was approximately five minutes.

Simulation Results

Objective results

Although the final measure of performance of any algorithm is the subjective quality of the processed picture, it is convenient to use a quantitative measure during the development phase of the algorithm. A commonly used quantitative (objective) measure for processed images is the peak signal to noise ratio (PSNR) and the mean square error (MSE) [28;29].

Table 2 shows the objective results of reconstructed pictures from different bands. The base band (HH) which has the most information has been used to reconstruct the original picture. Also, this band (HH) is combined with other bands to see objective and subjective quality. From Table 2, it can be concluded that GH band has the significant color information of the image. Also, it can be seen that the information in GG band can be neglected since the difference in the SNR between the reconstructed image from HH + GG bands and from HH band only is 0.0603 dB. The following section describes evaluation results using human subjects.

Subjective results

Table 3 represents the data collected on six images and Table 4 is a frequency table indicating the occurrence of each rating for each image. Figures 5 to 7 respresent black and white images used in testing the subjects, whereas, Figs. 8 to 10 are the colored images used to test the subjects. It also gives the standard error and 95 percent confidence intervals for mean. The analysis of variance results are presented in Table 5 indicating that there is a significant difference among the means *i.e.* the model is significant at a confidence level of 95% ($\alpha = 0.05$). Therefore, Ducan's multiple range test to compare individual means is conducted (see Table 6). The comparison indicates a significance between colored images 1 and 3 and black and white images 4 and 6 in the separate categories. Another factor which is evident from Duncan averages is that the colored image 1 is clearly superior to all other images tested. For the colored images, image 3 is the one closest to the conventional TV and image 1 is closest to the HDTV. This supports the objective measures.

Table 2. Simulation results show the reconstructed band(s)

Band(s) to be reconstructed	Number of samples to be send	Mean square error	Perceived signal to noise ratio (dB)	Has significant color information	Reconstructed image
HH	1/4 of the original samples	76.4405	29.2975	No	Image 4
HH+HG	1/2	63.4726	30.0981	No	Image 5
HH+GH	1/2	14.2977	36.5781	Yes	Image 2
HH+GG	1/2	75.3867	29.3578	No	Image 6
HH+HG+GH	3/4	1.4927	46.3709	Yes	Image 3
All bands	All the samples	0.4565	51.5366	Yes	Image 1

Table 3. Image ratings (colored and black and white images)

Image type	Colored			Black and white		
Image	1(Original)	2(1/2)	3(3/4)	4(1/4)	5(1/2)	6(1/2)
Total	149	137	126	111	128	140
Average	3.725	3.425	3.15	2.775	3.2	3.5

Table 4. Frequency of ratings for each image (colored and black and white images)

Image type	Colored			Black and white		
Rating frequency	1(Original)	2(1/2)	3(3/4)	4(1/4)	5(1/2)	6(1/2)
1	0	0	1	1	2	2
2	5	7	9	15	6	3
3	11	14	17	17	14	13
4	14	14	9	6	18	17
5	10	5	4	1	0	5

Fig. 5. Image 4 (reconstructed from 1/4 of the original samples).

Fig. 6. Image 5 (reconstructed from 1/2 of the original samples).

Table 5. One way anova (rating and image): Analysis of variance

Source of variation	Sum of squares	d.f.	Mean square	F-ratio	Sig. level
Between groups	21.771	5	4.345	4.989	.0002
Within groups	204.225	234	0.873		
Total (corrected)	225.996	239			

Table 6. Multiple range analysis for RATING by IMAGE

Method: Level	95% count	Duncan average	Homogeneous groups
IB4	40	2.775	*
IC3	40	3.150	**
IB5	40	3.200	**
IB6	40	3.500	**
IC1	40	3.725	*

Contrast	Difference
ICI-IC2	0.300
ICI-IC3	0.575 *
ICI-IB4	0.950 *
ICI-IB5	0.525 *
ICI-IB6	0.225
IC2-IC3	0.275
IC2-IB4	0.650 *
IC2-IB5	0.225 *
IC2-IB6	-0.075
IC3-IB4	0.375
IC3-IB5	-0.050
IB3-IB6	-0.350
IB4-IB5	-0.425
IB4-IB6	-0.725
IB5-IB6	-0.300

* denotes a statistically significant difference.

Another important finding is that only 5 out of 15 pairs (Table 6) were significantly different from one another indicating that subjects were not able to fully differentiate between certain images.

Discussion and Conclusions

It can be concluded that the quality of image can be detected by the human if it is significantly different from other images *i.e.* an improvement in the perception of the image can be realized if it is significantly different. The results of this study could be extrapolated to the case of the high resolution televison or HDTV where the resolution or the image quality is greatly enhanced and the human being can clearly distinguish between the conventional (low image quality) and HDTV (high image quality) with later being superior even on the subjective response. This is based on the fact that among the colored images, image 3 is the one closest to the conventional TV and image 1 is closest to the HDTV. This supports the objective measures. Hence it validates the improved design of HDTV over the conventional one by the response of human subjects. It is recommended that in the future work, images be presented on television (HDTV) instead of a projection screen so as to better evaluate the human subjects.

It can be seen from Table 6 that the subjects rated black and white image 6 better than image 5 which contradicts the objective quality PSNR. However, since the two images have the same bit rate and PSNR is also with 2% (Table 2) (PSNR for image 6 = 29.36 and for image 5 = 30.1), it was almost equiprobable to choose either image 5 or 6. The multiple range test on the subject responses indicates that the rating of images 5 and 6 are not significantly different from one another (Figs. 6 and 7) justifying the objective image data. Subjects clearly rated the black and white image 4 with 1/4 of the original samples and the color image 1, the full size image is the same as the worst and best respectively in the two groups. Colored images 2 and 3 were also confused by subjects indicating that they were not significantly different from one another. This finding would be justified by the fact that image 3 with 3/4 of the original samples with GG band in it did not provide additioal information to the subject. However, the two images have different bit rate and PSNR as indicated in Table 2. Image 3 is almost 10 dB greater than that for image 2. This will give us a good indication that the objective quality measure of images is not a reliable measure.

The findings of this study are of vital significance because it reinforces the objective claims that the high definition images used in high definition television are superior to the low definition images in conventional television as the colored image 1, which is closest to the HDTV image, was found to be significantly superior to other images. This supports the objective measures. Hence it validates the improved design of HDTV over the conventional one by the response of human subjects.

References

[1] Nelson, M.A. and Halberg, R.L. "Visual Contrast Sensitivity Function Obtained with Colored and a Chromatic Gratings." *Human Factors,* No. 21 (1979), 255-228.

[2] Steedman, W.C. and Baker, C.A. "Target Size and Visual Recognition." *Human Factors,* 2, No. 3 (1960), 120-127.

[3] Tannas, L.E. "HDTV Displays in Japan: Projection-CRT Systems on Top." *IEEE Spectrum,* 26, No. 10 (October, 1989), 31-33.

[4] Hopkins, R. "Digital HDTV Broadcasting." *IEEE Transactions on Broadcasting,* 37, No. 4 (December 1991), 123-127.

[5] Mckinney, J.C. "HDTV Approaches the End Game." *IEEE Transactions on Broadcasting,* 37, No. 4 (December 1991), 121-122.

[6] Nickelson, R.L. "The Evolution of HDVT in the Work of the CCIR." *IEEE Transactions on Broadcasting,* 35, No. 3 (September 1989), 250-258.

[7] Rubin, P.A. "Editorial on High Definition Television." *IEEE Transactions on Broadcasting,* 35, No. 3 (September, 1989), 249.

[8] Wassiczek, M.; Waters, G.T. and Wood, D. "European Perspective on HDTV Studio Production Standards." *IEEE Transactions on Broadcasting,* 35, No. 3 (September, 1989), 279-283.

[9] Keeler, E. "Inter Operability Considerations for Digital HDTV." *IEEE Transactions on Broadcasting,* 37, No. 4 (December 1991), 128-130.

[10] Zon, Y. "Digital HDTV Compression Techniques." *IEEE Transactions on Broadcasting,* 37, No. 4 (December, 1991),

[11] Zon, Y. "Comparison of Proposed Digital HDTV Terrestrial Broadcasting Systems." *IEEE Transactions on Broadcasting,* 37, No. 4 (December, 1991), 134-136.

[12] Lechner, B.J. "Testing HDTV Terrestrial Broadcasting Systems." *IEEE Transactions on Broadcasting,* 37, No. 4 (December 1991), 148-151.

[13] Jansky, D.M. "Methods for Accommodation of HDTV Terrestrial Broadcasting." *IEEE Transactions on Broadcasting,* 37, No. 4 (December 1991), 152-157.

[14] Hatori, M. and Yoshiro N. "1125/60 HDTV Studio Standard Intended to be a Worldwide Unified HDTV Standard." *IEEE Transactions on Broadcasting,* 35, No. 3 (September 1989), 270-278.

[15] Grether, W.F. and Baker, C.A. "Visual Presentation of Information." In: H.P. Vancott and R.C. Kindade (Eds.), *Human Engineering Guide to Equipment Design,* Washington, D.C.: US Government Printing Office, 1972.

[16] Erickson, R.A. "Live Criteria in Target Acquisition with Television." *Human Factors,* No. 20 (1978), 573-588.

[17] Williams, L.G. "The Effects of Target Specification on Objects Fined During Visual Search." *Perception and Psychophysics,* 1, (1966), 325-318.

[18] Gibson, J.J. "Cohat Gives Size to the Perception of Motion." *Psychological Review,* No. 75 (1968), 335-345.

[19] Gregory, R.L. *Eye and Brain.* 2nd Ed., New York: McGraw Hill, 1973.

[20] Christ, R.E. "Review and Analysis of Color Coding Research for Visual Displays." *Human Factors,* No. 17 (1975), 540-542.

Fig. 7. Image 6 (reconstructed from 1/2 of the original samples).

Fig. 8. Image 2 (reconstructed from 1/2 of the original samples).

Fig. 9. Image 3 (reconstructed from 3/4 of the original samples).

Fig. 10. Image 6 (reconstructed from all the original samples).

images. This supports the objective measures. Hence it validates the improved design of HDTV over the conventional one by the response of human subjects.

References

[1] Nelson, M.A. and Halberg, R.L. "Visual Contrast Sensitivity Function Obtained with Colored and a Chromatic Gratings." *Human Factors*, No. 21 (1979), 255-228.

[2] Steedman, W.C. and Baker, C.A. "Target Size and Visual Recognition." *Human Factors*, 2, No. 3 (1960), 120-127.

[3] Tannas, L.E. "HDTV Displays in Japan: Projection-CRT Systems on Top." *IEEE Spectrum*, 26, No. 10 (October, 1989), 31-33.

[4] Hopkins, R. "Digital HDTV Broadcasting." *IEEE Transactions on Broadcasting*, 37, No. 4 (December 1991), 123-127.

[5] Mckinney, J.C. "HDTV Approaches the End Game." *IEEE Transactions on Broadcasting*, 37, No. 4 (December 1991), 121-122.

[6] Nickelson, R.L. "The Evolution of HDVT in the Work of the CCIR." *IEEE Transactions on Broadcasting*, 35, No. 3 (September 1989), 250-258.

[7] Rubin, P.A. "Editorial on High Definition Television." *IEEE Transactions on Broadcasting*, 35, No. 3 (September, 1989), 249.

[8] Wassiczek, M.; Waters, G.T. and Wood, D. "European Perspective on HDTV Studio Production Standards." *IEEE Transactions on Broadcasting*, 35, No. 3 (September, 1989), 279-283.

[9] Keeler, E. "Inter Operability Considerations for Digital HDTV." *IEEE Transactions on Broadcasting*, 37, No. 4 (December 1991), 128-130.

[10] Zon, Y. "Digital HDTV Compression Techniques." *IEEE Transactions on Broadcasting*, 37, No. 4 (December, 1991),

[11] Zon, Y. "Comparison of Proposed Digital HDTV Terrestrial Broadcasting Systems." *IEEE Transactions on Broadcasting*, 37, No. 4 (December, 1991), 134-136.

[12] Lechner, B.J. "Testing HDTV Terrestrial Broadcasting Systems." *IEEE Transactions on Broadcasting*, 37, No. 4 (December 1991), 148-151.

[13] Jansky, D.M. "Methods for Accommodation of HDTV Terrestrial Broadcasting." *IEEE Transactions on Broadcasting*, 37, No. 4 (December 1991), 152-157.

[14] Hatori, M. and Yoshiro N. "1125/60 HDTV Studio Standard Intended to be a Worldwide Unified HDTV Standard." *IEEE Transactions on Broadcasting*, 35, No. 3 (September 1989), 270-278.

[15] Grether, W.F. and Baker, C.A. "Visual Presentation of Information." In: H.P. Vancott and R.C. Kindade (Eds.), *Human Engineering Guide to Equipment Design*, Washington, D.C.: US Government Printing Office, 1972.

[16] Erickson, R.A. "Live Criteria in Target Acquisition with Television." *Human Factors*, No. 20 (1978), 573-588.

[17] Williams, L.G. "The Effects of Target Specification on Objects Fined During Visual Search." *Perception and Psychophysics*, 1, (1966), 325-318.

[18] Gibson, J.J. "Cohat Gives Size to the Perception of Motion." *Psychological Review*, No. 75 (1968), 335-345.

[19] Gregory, R.L. *Eye and Brain*. 2nd Ed., New York: McGraw Hill, 1973.

[20] Christ, R.E. "Review and Analysis of Color Coding Research for Visual Displays." *Human Factors*, No. 17 (1975), 540-542. .

[21] Kellog, R.S.; Kennedy, R.S. and Woodruff, R.R. "A Comparison of Color Versus Black & White Visual Display as Indicated by Bombing Performance in the 2835 TA-45 Fight Simulator." *Proceedings of the Human Factors Society,* 25 (1981), 233-234.

[22] Rogowitz, B.E. Displays: The Human Factor." *Byte,* 7, No. 17 (July 1992), 195-200.

[23] Esteban, D. and Galand, C. "Application of Quadrature Mirror Filters to Split Band Voice Coding Schemes." *Proc. ICASSP,* (May, 1977), 191-195.

[24] Gharavi, H. and Tabatabai, A. "Sub-band Coding of Digital Images Using Two-Dimensional Quadrature Mirror Filtering." *Proc. SPIE,* 707 (September 1986), 51-61.

[25] Johnson, J.D. "A Filter Family Designed for Use in Quadrature Mirror Filter Banks." *Proc. ICASSP,* (April 1980), 291-294.

[26] Vetterli, M. "Multi-dimensional Sub-Band Coding: Some Theory and Algorithms." *Signal Processing,* 6 (April, 1984), 97-112.

[27] Woods, J.W. and O'Neil, S.C. "Sub-Band Coding of Images." *IEEE Trans. Acoust., Speech. Signal Processing,* ASSP-34 (October, 1986), 1278-1288.

[28] Al-Asmari, A.K. "Sub-band Coding of HDTV Images Using GQMF." *IEE Electron Lett.,* 28, No. 14 (July 1992), 1335-1337.

[29] Coppisetti, N.; Kwatra, S.C. and Al-Asmari, A.K. "Low Complexity Sub-band Encoding for HDTV Images." *IEEE J. Select Areas Commun.,* 11, No. 1 (1993), 77-78.

قياس الجودة البصري لنظام الصور الرقمية

عوض خزيم الأسمري* و سيد عابد نقوي**

** قسم الهندسة الكهربائية و ** قسم الهندسة الميكانيكية، كلية الهندسة، جامعة الملك سعود،*

ص. ب. ٨٠٠، الرياض ١١٤٢١، المملكة العربية السعودية

(سُلِّم في ١٣/١٢/١٩٩٤م؛ قُبِل للنشر في ١٤/٦/١٩٩٥م)

ملخص البحث. تقدم هذه الورقة دراسة ميدانية لتمييز ٦ صور رقمية مختلفة (٣ملونة، ٣ أبيض وأسود) على عيّنة شملت ٤٠ شخصًا باستخدام مقياس الجودة البصري. وتضم الورقة أيضًا مقارنة للنتائج المستخلصة باستخدام التحليل التبايني واختبار المقارنة الزوجي. وقد وجد أن استجابة الأشخاص ترجح للصور الملونة ذات الاستبانة العالية. وهذا يثبت صلاحية الافتراض القائل بأن: الصور الرقميّة ذات الاستبانة العالية أفضل من الصور الرقمية المعتادة. وعند مقارنة كل زوج من الصور وجد أن الاختلاف لكل ٥ من ١٥ واضح وكذلك من صورة لأخرى. وهذا يؤكد أن الصور ذات الاستبانة المتقاربة لا يمكن مفاضلتها من قبل الأشخاص. وتعطي هذه الدراسة مدلولاً حيويًا لأنها تؤكد عمليًا صحة الإدعاء الموضوعي بأن الصور ذات الاستبانة العالية والمستخدمة في التليفزيون ذي الاستبانة العالية (HDTV) أفضل من الصورة ذات الاستبانة المنخفضة والمستخدمة في التليفزيون العادي.