DETAIL SELECTION INCORPORATING SUBJECTIVE FACTORS FOR VERY LOW BIT-RATE IMAGE CODING

Yongqin Zeng

Signal Processing and Digital Systems Section,
Department of Electrical and Electronic Engineering,
Imperial College of Science, Technology and Medicine, London SW7 2BT, UK.
y.zeng@ic.ac.uk

ABSTRACT

This paper is concerned with the subjective selection of image details for segmentation-based very low bit-rate coding. An approach is proposed for extracting image details and ranking the details in the order of perceptual significance to the Human Visual System (HVS). A perceptual ranking model has been established by multivariate regression analysis based on the data obtained from subjective tests. In terms of the correlation between the objective ranking and subjective ranking of perceptual significance and subjective image quality, this model provides better results in detail ranking than that provided by an empirical ranking formula proposed in a previous study. Segmented regions and selected details are coded to illustrate the efficiency of coding meaningful details. Compared with the pure segmented images, the addition of the selected details improves the subjective image quality at lower bit rates.

Key words: detail, HVS, image coding, segmentation-based, very low bit-rate.

1. INTRODUCTION

The popular segmentation-based coding schemes are based on contour/texture representation image model [1]. The major advantage of these approaches is that there is no requirement for any prior knowledge or assumptions about specific objects in the scene and therefore they can be applied to general classes of images. However, the principal difficulty in segmentation-based approaches is that segmentation algorithms lack the ability to appropriately segment the image into regions consistent with the areas perceived by the human eye. When the number of segmented regions increases, more undesirable contours are generated which reduce the coding efficiency. Therefore, a small number of segmented regions are usually coded in very low bit-rate schemes to achieve a high compression ratio, whereas textures are coded by waveform coding techniques. Consequently, details of an image which remain in the texture are usually lost in the texture coding process due to the lacking of homogeneity and periodicity. Nevertheless, some of the details are very significant from the visual point of view. The absence of these details could obviously degrade the subjective image quality.

However, the quantity of extracted details is usually quite large and the details are not all of equal interest for human observers. Moreover, the coding of details are bit-consuming. Therefore, the visual significance of each detail is of great importance in selecting details to be coded. The precision of detail ranking plays a key role in improving the visual quality of a segmented image.

A robust ranking algorithm will rely primarily on the progress in understanding the mechanism of human vision. Due to the lack of related knowledge, a widely-agreed model of perceptual ranking has not been reported in the literature. An attempt was made previously to extract these details by using morphological operators and to select perceptual significant details [2, 3]. A preliminary approximation of perceptual ranking has been performed by an empirical formula. This approach with perceptual detail ranking has proved both valid and effective in improving image quality at a very low bit rate. However, neither the mathematical basis nor the experimental procedure of obtaining this formula has been discussed. In addition the accuracy of perceptual ranking needs to be further estimated by providing quantitative information on the correlation between perceptual ranking and subjective evaluation.

In this study, a modified detail extractor was proposed [4] and efforts were made to provide a statistical model of perceptual ranking based on the data from the subjective test of extracted details. The proposed approach was applied in a segmentation-based coding scheme. The results were compared with those produced by the empirical formula, by the pure segmentation-based coding scheme and by the JPEG standard.

The paper is organised as follows. Section 2 describes the procedure of establishing the perceptual ranking model. Results are shown in section 3. Conclusions are presented in section 4.

2. DETAIL RANKING AND SELECTION

The crucial steps to establish the ranking model are: selecting variables, data collection and finding the weight of each variable for the designed model.

2.1. Perceptual variables

By taking into account the visual mechanism, *contrast, area, shape* and *activities of background around each detail* are believed to be the most important variables in describing the details extracted from *grey level* and *still* images. In this approach, the contrast is calculated as the mean value of grey level of M pixels in each detail i:

$$Contrast(i) = \frac{1}{M} \sum_{j=1}^{M} x_j \qquad (1)$$

The area of a given detail is described by the number of pixels within each detail; The shape of a detail is described by the form-factor [5]:

$$Formfactor(i) = \frac{4 * \Pi * Area}{Perimeter^2} \qquad (2)$$

The background activity is calculated by the standard deviation of the surrounding background of each detail:

$$Bdeviation(i) = \frac{1}{M} \sum_{j=1}^{M} |x_j - \overline{x_i}| \qquad (3)$$

$$\overline{x_i} = \frac{1}{M} \sum_{j=1}^{M} x_j \qquad (4)$$

where M is the area of surrounding background. The extracted details are labelled individually to be described by the selected perceptual variables.

2.2. Data collection for modelling

In order to establish a ranking model, training data and test data are needed. The training data were collected from one set, 173 details included in one image, while the test data were obtained from three sets, 239 details in other three images which were used to evaluate the reliability of the obtained model.

2.2.1. Dependent data: subjective test

The dependent data were the subjective scores of the details. The significance of each detail corresponding to the human perception can only be measured by subjective testing. Since there is no existing method for subjective detail ranking, a seven level ranking method is designed for examining the effect of supplementing the segmented image with each detail. The opinion rating scale is shown in Table 1, which is evolved from the *Subjective Impairment Scale* by K. R. Rao [6].

The test was carried out under the assumption of the low-level attention mechanism. In other words, each detail was judged by every observer who was neither task-oriented nor knowledge-based. It was still impossible to avoid the unconscious *high-level* judgement depending on the experience of individual observers. The statistical mean value of detail significance among all observers can compensate the bias to some extent. The experiment was implemented as described below: (1) Five image coding researchers and six non-professional people with normal visual acuity attended the experiment under a uniform experimental condition; (2) Each test image is generated from a segmented image overlapped by one of the extracted details. Four sequences of test images were produced by overlapping four sets 412 details extracted from corresponding residues of their segmented images.

Opinion	*Score*
Extremely effective in improving the image quality	7
Definitely effective in improving the image quality	6
Somewhat effective in improving the image quality	5
Improvement to the image quality but not effective	4
Definitely noticeable but only slight improvement	3
Barely noticeable	2
Not noticeable	1

Table 1: The opinion rating scale

In order to identify the similarity of subjective judgements from different people, data correlation was analysed to indicate the relationship between every two sets of scores. It is shown that the scores are significantly related to each other within the confidence level of 0.05. The results of the subjective classification from eleven subjects were presented by using the Mean Opinion Score (MOS):

$$MOS = \frac{\sum_{i=1}^{K} C_i}{K} \qquad (5)$$

where C_i is the numerical value corresponding to category and K is the number of subjects involved.

2.2.2. Independent data

The independent data, including contrast, area, texture and form-factors of each detail, were obtained by an automatic measure algorithm.

2.3. Regression analysis of perceptual variables

Multivariate techniques provide the mathematical methods to investigate the statistic relationship between the dependent and explanatory variables. Considering the purpose of this study, multiple regression analysis [7] was employed in this modelling.

2.3.1. Multivariate regression model

Because the visual mechanism is too complex to fully understand and the nonlinear characteristics of the selected variables have also not been described quantitatively, the most practical model of perceptual ranking is considered as:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \varepsilon \qquad (6)$$

where X_1, X_2, X_3, X_4 represent the variables of contrast, area, background activity and form-factor respectively; $\beta_0, \beta_1, \beta_2, \beta_2, \beta_4$ are the regression coefficients that need to be estimated; the Y is an observable random variable; and the ε is the error component. The estimated values of coefficients were then obtained through the *least-squares approach.*

The backward elimination procedure was used for selecting the variables and producing the estimated coefficients. In order to evaluate the ranking model, four criteria

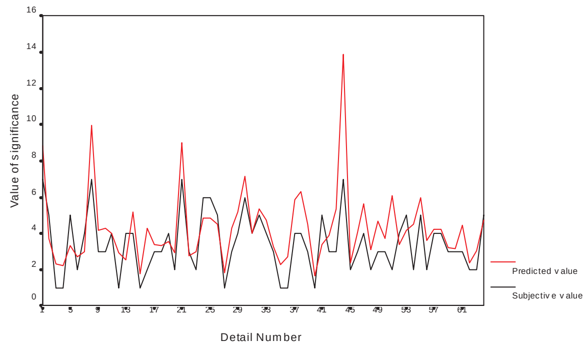


Figure 1: Bright and dark extracted details of *salesman* and the line graphs of predicted and observed values

were considered: sample squared multiple correlation R^2; F test value; Durbin-Watson value; and autocorrelation functions (ACF) of the residual series. A statistical software package called *Statistical Package of Social Science* (SPSS) was utilised to perform the regression and produce an equation for predicting detail significance.

There were no variables removed from the selected model by the backward elimination procedure. The fitted prediction equation, therefore, is:

$$\begin{aligned}\hat{\mathbf{Y}} &= \mathbf{1.28 + 0.06 * Contrast + 0.02 * Area -} \\ &\quad \mathbf{0.03 * Texture - 0.57 * Formfactor} \qquad (7)\end{aligned}$$

The R^2 value (0.76) and the F ratio (145.16) indicate that the obtained model is sound. The Durbin-Watson value (1.70) shows that there is some positive correlation in the residual series. It means that there are still some small patterns retained in the residual. Seventy-six percent variation of Y can be explained by independent variable X's.

2.3.2. Evaluating the reliability

The model is reliable only if the obtained model predicts well for the subsequent samples. The comparison of the subjective and the predicted values of details from one of the test images is given in Figure 1. It is shown that the majority of variation of subjective values are caught by the predicted values, particular at the high level of importance.

The cross-validation analysis [7] was also tested using the fitted model. The shrinkage of the three test sets of detail images are small, which indicates that the model is reasonably reliable. As a comparison, the squares of correlation coefficients between the values predicted by the cited formula [3] and the subjective scores were also calculated. It suggests that the fitted model shows higher reliability than the cited formula.

2.4. Selected details

How many details should be selected for coding depends on the distribution of the details and the available bite rate. According to the histograms of ranked details, the details in level 7 and level 6 are within the top 15% of the ranked details. In addition, 15% of details is acceptable to maintain available bit rate. Therefore, the top 15% of the details are selected for coding. Nevertheless, it should be noted that in the head and shoulder picture the number of extracted details is relatively small but the majority of the details are usually important. Hence, the minimum number of selected details is suggested to be not less than 10 for coding.

3. RESULTS AND DISCUSSIONS

The proposed approach was applied in a segmentation-based image coding scheme. Firstly, the original image (Figure 2(a)) was segmented by the RSST segmentation algorithm [8]. The details were then extracted from the residue of the segmented image (Figure 2(b)) [4]. The top 15% of extracted details were selected by the established ranking model. Finally, both of the segmented regions and the selected details were coded by chain coding algorithm [9] to illustrate the efficiency of coding meaningful details. Figure 3 provides the results produced by different approaches. Figure 3(b) shows good subjective quality but costs a rather high bit-rate. In Figure 3(d), some significant details were included in the selected detail images, but some details, such as the details on the faces, were lost. These might be caused by the over-weighted texture parameter [3]. By contrast, the image compensated with the same number of details selected by the established model (Figure 3(f)) shows better subjective quality than the image of Figure 3(d). Meanwhile, compared with the image merely segmented in more regions (Figure 3(g)) (0.17 bpp), the image compensated by meaningful details shows better subjective image quality by using a lower bit-rate (0.12 bpp). This shows that the selected details provide the effective information from human visual point of view. Finally, the comparison between the image coded by the proposed approach (Figure 3(f)) and the image coded by the JPEG (Figure 3(h)) indicates the advantages of taking into account of the perceptual factors in coding at a very low bit-rate.

4. CONCLUSIONS

This work has been concentrated on perceptual ranking and selection of image details. A perceptual ranking model was established through multivariate regression analysis. This model has provided good results in detail selection in terms of both subjective image quality and the correlation between the objective ranking and subjective ranking. Compared with the pure segmentation-based coding and JPEG, the coding of both segmented regions and details shows better subjective image quality at a lower bit rate. The false contours corresponding to variations within homogeneous region are avoided due to coarse segmentation, while the false contours corresponding to less meaningful details are removed by detail selection. This study improved the ability of the image encoder to distinguish the desirable information from undesirable information for coding at a very low bit rate. It is suggested to evolve the ranking model by establishing a nonlinear model using neuron network technology.

(a) Original image

(b) Segmented image in 10 regions, 0.06 bpp

Figure 2: Original image of "Lenna" and the image segmented in 10 regions.

5. ACKNOWLEDGEMENT

The author would like to thank Professor A. G. Constantinides, Imperial College, Science, Technology and Medicine, for his supervision during this study.

6. REFERENCES

[1] M. Kunt, A. Ikonomopoulos, and M. Kocher. Second generation image coding techniques. *Proc. IEEE*, 73(4):549–574, Apr 1985.

[2] J. R. Casas and L. Torres. Coding of details in very-low bit-rate video system. *IEEE Transaction on Circuits and Systems for Video Technology*, 4(3):317–327, June 1994.

[3] J. R. Casas and L. Torris. Coding of significant features in very low bit-rate video systems. In *SPIE Visual Communications and Image processing'94*, 1994.

[4] Y. Zeng. Region coding and detail compensation. In X. Tang B. Yuan, editor, *International conference on signal processing proceedings*, volume 2, pages 1227–1230, Beijing, China, October 1996.

[5] J. C. Russ. *Computer-Assisted Microscopy the measurement and analysis of Images.* Plenum Press, 1990.

[6] K. R. Rao. *Discrete Cosine Transform: algorithms, advantages, applications.* Academic press, 1990.

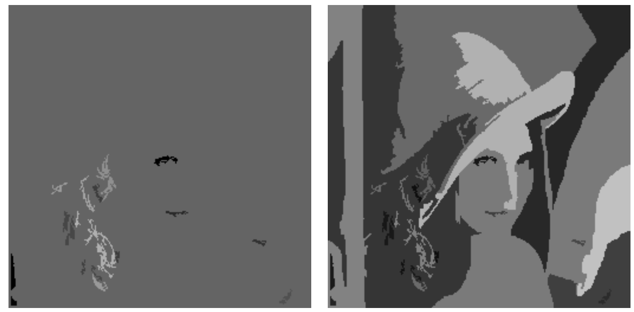
[7] Kleinbaum and Muller Kupper. *Applied Regression Analysis and Other Multivariable Methods.* PWS-KENT Publishing Company, 1988.

[8] O. J. Morris, M. de J. Lee, and A. G. Constantinides. Graph theory for image analysis: An approach based on the shortest spanning tree. *Proc. IEE*, 133 Part F:146–152, Apr. 1986.

[9] M. J. Biggar. Thin line coding techniques. In *Proc. Int. Conf. DSP*, pages 471–475, Florence, September 1987.

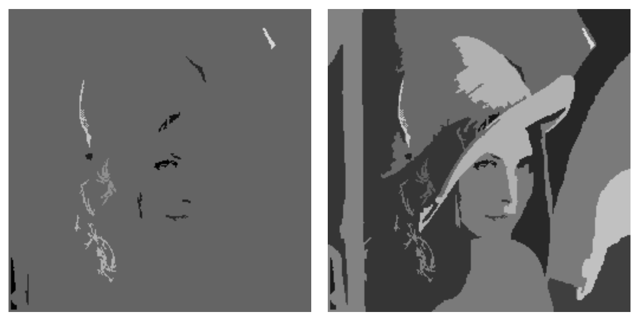
(a) All extracted details, 0.21 bpp

(b) 10 regions segmented image with all extracted details of (d), 0.31 bpp



(c) 20 details selected by empirical formula, 0.05 bpp

(d) 10 regions segmented image with 20 details of (c), 0.12 bpp



(e) 20 details selected by established model, 0.05 bpp

(f) 10 regions segmented image with 20 details of (e), 0.12 bpp

(g) Segmented image in 50 regions, 0.17 bpp

(h) Image coded by JPEG, 0.18 bpp

Figure 3: The comparison of the results from different approaches.