Image Compression and Reconstruction Using a 1-D Feature Catalogue

Brian Y.K. Aw[1]*, Robyn A. Owens*[1] *and John Ross*[2]
[1] Department of Computer Science, University of Western Australia.
[2] Department of Psychology, University of Western Australia.

Abstract. This paper presents a method of compressing and reconstructing a real image using its feature map and a feature catalogue that conprises of feature templates representing the local forms of features found in a number of natural images. Unlike most context-texture based techniques that assume all feature profiles at feature points to be some form of graded steps, this method is able to restore the shading in the neighbourhood of a feature point close to its original values, whilst maintaining high compression ratios of around 20:1.

1. Introduction

The image compression ratio achieved by early image coding techniques based on information theory operating on natural images saturated at a value of 10:1 in the early eighties [KIK1]. Later techniques that code an image in terms of its feature map managed to obtain higher compression ratios but at a sacrifice of image quality, namely the loss of original local luminance form (i.e. feature profiles) at feature points. The technique described in this paper is able to correct these defects yet maintains compression ratios around 20:1.

An image can be decomposed into two parts: a feature map and a featureless portion. In other existing techniques (see review article [KIK1]), the feature map is thresholded and only the location of feature points is coded. Consequently, all information about the original luminance profiles that give rise to those feature points is lost in the reconstruction phase, where artificial graded step profiles are used instead. To recover this lost information, our technique makes use of a common feature catalogue that consists of a number of 1-dimensional feature templates. Each template describes a feature profile in terms of normalised mean luminance values and standard deviations at various pixel locations of the profile. In [AOR1], it has been shown that a catalogue whose templates approximate closely most feature luminance profiles in many natural images can be derived from some appropriate sample images. In the coding phase of our technique, both the locations of feature points and pointers indicating feature types are also retained. Each pointer points to the feature templates in the catalogue that best approximates the original luminance profile in the neighbourhood of the indexing feature point. Subsequently, the information encoded in the pointers is used to recover the luminance profile of features at the various locations in the image. The same feature catalogue is used for all images. In our technique, we also encode the featureless portion of an image in terms of a small number of fourier coefficients which is used in a later stage to recover the background shading in the reconstructed image.

2. The feature catalogue

The feature catalogue mentioned above is shown in Fig. 1. This feature catalogue is formed by a 2-layer recurrent neural network guided by the local energy operator. Please refer to [AOR1] and [MO1] for further details on the network and the local energy model. The horizontal axis is the spatial dimension in pixel units (1 to 5), and the vertical axis represents luminance values from 0 (lowest point) to 255 (highest point). The feature is located at pixel 3. The mean luminance values of each template are marked by horizontal white bars. A gaussian function is plotted as shades along the vertical axis for each pixel location of a feature profile. The wider the spread of the shading, the larger the standard deviation value.

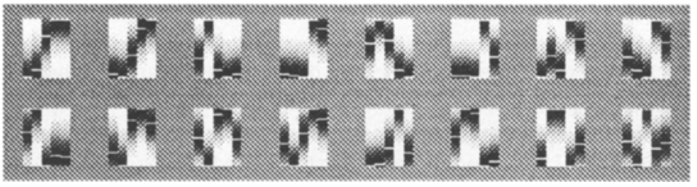

Figure 1. Feature Catalogue.

3. The coding process

An image to be compressed is coded in portions: the feature regions and the featureless regions. The feature regions consist of feature points (*i.e.* pixel locations) and the luminance profiles in the neighbourhood of the points (*i.e.* feature profiles). The feature points are defined at the peak of the local energy according to the local energy model. The original feature profile is rescaled (into x_i) such that it has least square errors with the respective feature template (μ_i and σ_i) in the common catalogue. The similarity index (z) for the comparison between the scaled feature profile (x_i) and a feature template is defined as follows:

$$z = \frac{1}{N} \sum_{i=1}^{N} \exp\left[- \frac{(x_i - \mu_i)^2}{\sigma_i^2} \right] \tag{1}$$

where N (=5) is the number of pixels in the 1-D templates.

The template that produces the highest z value is used to represent the feature profile at the feature point. A pointer from this feature point is then set up and coded with the best-matched template number and the necessary rescaling parameters (a d.c. shift and a multiplier).

For example, the feature point map of an original image "Baby" (Fig. 2(a)) is shown in Fig. 2(b). This map combines the feature points found in the horizontal and vertical directions. Centred at each (black dot) location, the scaled 1-D profile of the best matched feature type from the catalogue is shown in Fig. 2(c). Some features are in the horizontal direction and some are in the vertical direction. Visually, it is evident that the information retained in Fig. 2(c), *i.e.* location plus local form, is richer than just the locational information itself represented in Fig. 2(b).

Besides coding the feature portion of an image, we also code a low-pass version of its featureless portion in terms of the coefficients of its low frequency harmonics. For a 256×256-pixel image, we retain the lowest 10×10 2-dimensional complex coefficients of its FFT (Fig. 2(d)).

We can attain a compression ratio of 20:1 if we (a) assume around 2% of the original image pixels are feature points either in the horizontal and/or vertical direction, and (b) use the following code words for the various messages of an image:

- an average of 4.5 bits per feature point to code the positional information in Huffman code;
- a 5-bit code word to code the d.c. parameter.
- a 4-bit code word for the multipling (scaling) parameter.
- a 3-bit code word to code the feature template number;
- a 1-bit code word to indicate the 1-D feature direction (horizontal or vertical);
- a 16-bit code word for the complex FFT coefficients.

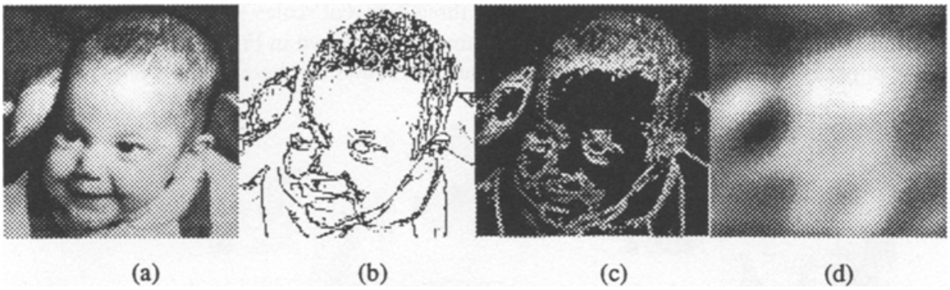

Figure 2. (a) Original image "Baby". (b) The locations of feature points (shown as black dots). (c) The template profile of the best matched feature template from the catalogue for each feature point in the image is superimposed at the location of that point. (d) The low-passed version of "Baby". Only the lowest 10×10 2-dimensional complex coefficients of the FFT of "Baby" are retained.

4. Results of reconstruction process

There are two stages involved in the reconstruction process. First, luminance profiles at feature points are retrieved from the catalogue of feature templates by means of the coded pointers. Second, an iterative process toggles between local averaging and fourier coefficient foldback. Local averaging smooths unwanted artifacts in featureless regions but alters the original low frequency harmonics of an image. The coded fourier coefficients are used to reinforce these harmonics at each iteration.

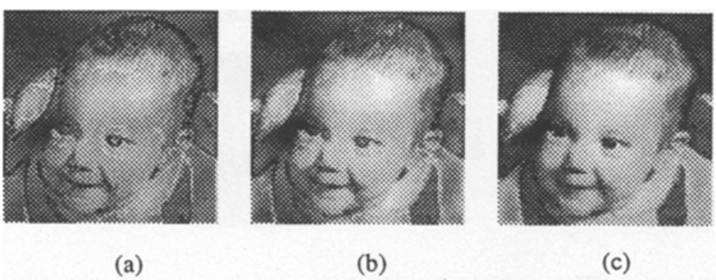

Figure 3. (a) Initial stage of reconstruction of the image "Baby". Feature profiles and low-pass FFT coefficients are retrieved from the compressed data. (b) Result after 10 iterations. (c) Result after 50 iterations.

The 2-dimensional local averaging process takes place in local neighbourhoods of sizes that depend on the location of the current point in the featureless regions. The nearer the current point is to a feature location, the smaller the size of the averaging neighbourhood. Only smoothness up to the first order is enforced by the averging action. An illustration of the reconstruction process is shown in Fig. 3. The first image (a) is the initial result when the coded feature local forms and the low-pass data are retrieved. This is followed by the local averaging operation interlaced by the reinforcement of the coded low-pass fourier coefficients. The results after 10 and 50 rounds of iterations are shown in (b) and (c) respectively.

5. More experimental results

The coding and reconstruction scheme proposed in this paper is used to test three other natural images: "Machine", "Animal" and "X-Ray", all at three different scales (Fig. 4). If L_n denotes the scale of an image of size $n \times n \times 8$-bit, then the three scales shown in Fig. 4 are L_{256}, L_{128} and L_{64}. The upper rows are the original images and the lower rows the reconstructed ones.

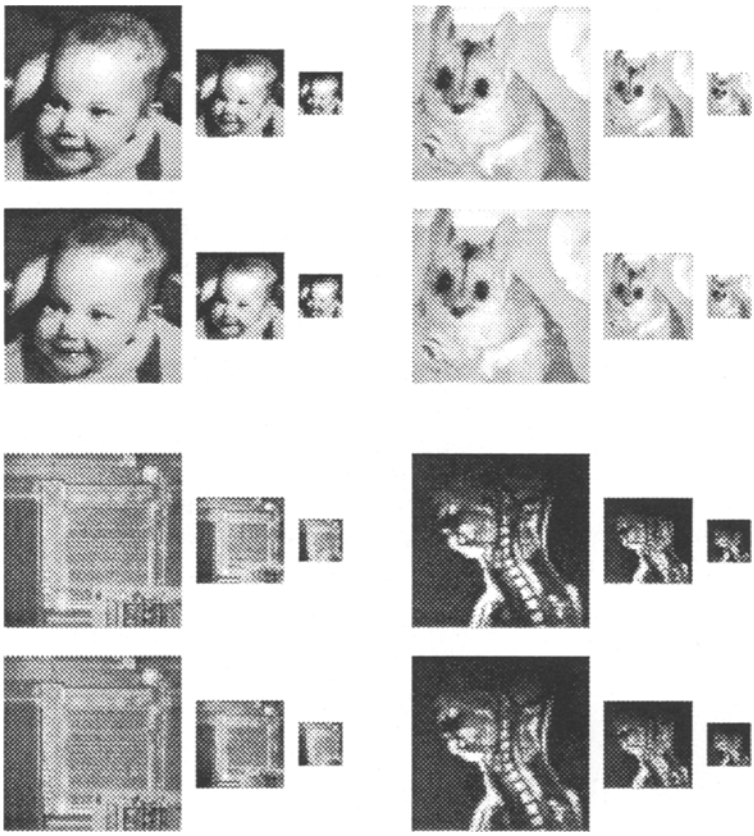

Figure 4. Test images. The original images at different scales are in the upper rows and the reconstructed ones are in the lower rows. The images are named "Baby" (top-left), "Animal" (top-right), "Machine" (bottom-left) and "X-Ray" (bottom-right). The three different image sizes are 256×256, 128×128 and 64×64 pixels.

An error percentage per pixel on the basis of a 256 level grey scale is computed (Table 1) as an indication of the difference between the original image and its reconstructed version. The error is calculated as the average of the absolute difference in luminance values between corresponding pixel locations in the two images.

Image size	"Baby"	"Animal"	"X-Ray"	"Machine"
256×256	1.6	1.1	1.5	1.4
128×128	1.6	1.0	1.4	1.8
64×64	1.5	1.1	1.7	2.1

Table 1. Average percentage difference per pixel between the original and reconstructed images.

6. Conclusions

It was noted that recent image compression and reconstruction technique achieved high compression ratios at the expense of losing information at feature points. This is due to the fact that a more fundamental question was yet to find its solution, that is: What is a feature? In a previous paper [AOR1], it is shown that features can assume a wide variety of local luminance forms. In our analysis, a catalogue of 16 templates was sufficient to accommodate most feature profiles encountered in a number of natural images. By setting up pointers to this catalogue, we have shown that the local luminance forms of features can be preserved in a compressed format. Subsequently, local feature forms are reproduced faithfully in the reconstruction process.

It is also demonstrated that a 2-dimensional averaging algorithm, bolstered by the incorporation of the FFT coefficients of lower harmonics of the original image, is able to reconstruct a good quality image from its compressed feature map.

In terms of efficiency, this scheme achieves a compression ratio of around 20:1 with little sacrifice in image quality. A higher compression ratio is within reach by incorporating more efficient techniques in the encoding of feature maps. One possible way is to first obtain a closed contour by using standard region-growing methods in the local energy maps.

It might also be possible to make further gains by taking advantage of the fact that similar features recur at different scales (see also [AOR1]). Reconstruction might be possible in stages, beginning at the coarest scale and then moving up the scale pyramid, introducing at finer scales only features that first occur at those scales. The scheme of transmitting image data in stages according to scales of interest makes good sense in terms of the use of transmitting media.

References

[AOR1] Aw, B.Y.K., Owens, R.A., Ross, J.: A catalogue of 1-d features in natural images. Manuscript submitted to Neural Computation. (1991)

[KIK1] Kunt, M., Ikonomopoulos, A., Kocher, M.: Second-generation image-coding techniques. Proc. of the IEEE. **73** (1985) 549-574

[MO1] Morrone, M.C., Owens, R.A.: Feature detection from local energy. Pat. Recog. Letters. **6** (1987) 303-313.