Stripwise Image Warping

L. Baghai-Ravary[‡], S. W. Beet[†] and M. O. Tokhi[‡]

[‡] Department of Automatic Control and Systems Engineering,
[†] Department of Electronic and Electrical Engineering,
Sheffield University, Mappin Street, Sheffield, S1 3JD, UK.

Abstract

This paper presents a new approach to image coding.

An image is segmented into clusters of similar feature densities using an automatic segmentation algorithm. Each cluster takes the form of a whole strip across the image. These are then treated as series' of vectors, some of which may be omitted as they can be interpolated from other vectors within their cluster. This gives a uniform feature density and consistent statistics across the whole image in addition to encoding and compressing the image data.

Examples are given to illustrate the degree of compression attainable, the nature of the distortion introduced and the advantages of adopting a strip-wise approach.

Introduction

A common problem with conventional image coding techniques is an inconsistency of the image statistics, due to the non-uniform feature density within the image. This makes it difficult to design a single vector quantisation codebook appropriate for all parts of the image. However, this problem could be ameliorated if a sophisticated warping function were applied, which balanced out the feature density and the image statistics.

Defining an arbitrary warping function would require an excessive and unrealistic level of side information. It is simpler just to remove series of rows or columns of pixels which may be interpolated from their neighbours [1]. This requires much less side information to identify the omitted data vectors.

Unfortunately, it is unusual for an image to have many whole rows or columns of pixels which could be estimated from the rest of the image, unless the picture is very simple or sparsely populated. However, if the image were segmented, many short rows or columns could be estimated from other parts of the image in the same segment.

The value of each pixel normally changes smoothly with respect to its neighbours and could be interpolated from them. Thus, if the image were initially segmented into non-uniform-width strips, each consisting of similar feature densities, these smaller segments would have a structure which makes them highly predictable. Thus many of the vectors which make up each strip can be omitted, and the image effectively compressed.

One technique which can be used to identify predictable vectors, and to regenerate them during reconstruction of the image, is multi-step adaptive flux interpolation (MAFI). This is a non-linear method for identifying redundant vectors within a series. This algorithm can, therefore, be used to group similar vectors, and to eliminate them, for segmentation or coding purposes respectively.

Figure 1 shows the block diagram for the experimental procedure of this paper.

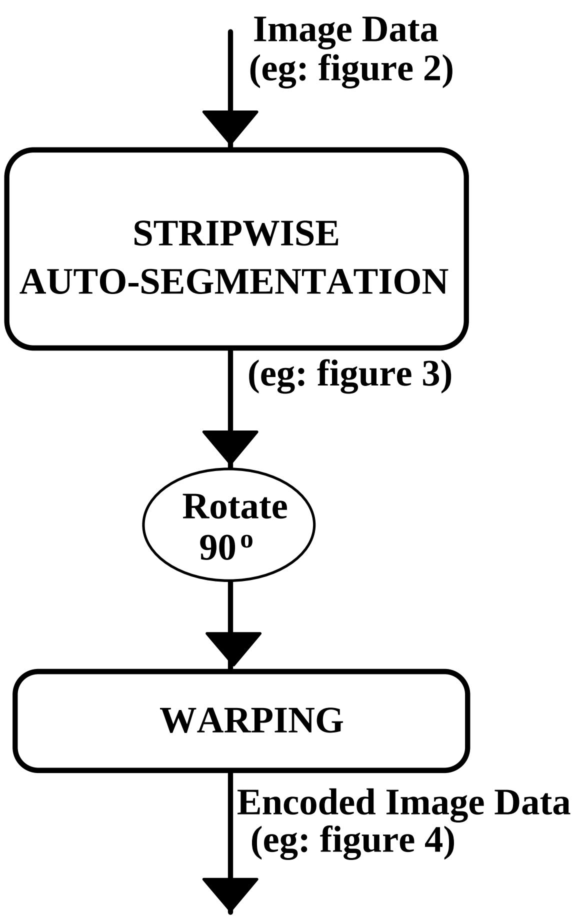

Figure 1: Block diagram of the image coding procedure.

The encoded image data is composed of several strips, with different widths, which have been warped and reduced in length.

Automatic segmentation

The first step in this method of image compression is to divide the image into clusters which consist of a number of vectors forming a strip of the image. The length of vectors in the strip (or the strip width) is automatically chosen to be of an appropriate size to maintain a smooth flow of features, at least in the transverse dimension.

While it would be possible to make these strips of a fixed width, defined *a priori*, a smaller number of strips can be used if the strip boundaries are chosen to coincide with the edges of the dominant features in the image (i.e. those which are largest and with the greatest contrast). These can be identified by performing a coarse MAFI analysis [1] on the image to identify groups of similar, say, columns.

Each of these column groups is then treated as a strip and compressed by removing redundant transverse vectors with a fine MAFI analysis, applied vertically to remove horizontal data vectors.

Figure 2 shows an image of a clown taken from MATLAB. This image consists of a mixture of low and high density of features. In particular, the clown’s fine, curly hair consists of a detailed, but visually structured, pattern which would be a a difficult task for any image coder to compress.

Figure 3 is an example of the automatically-determined strip boundaries, identified by MAFI, using the image data in figure 2. It is clear that, in the main, these coincide with visually significant features in the image. Examining each strip also reveals that a very large proportion of the horizontal lines (vectors) are highly correlated with their neighbours (i.e. many of them are redundant). On perusal of each strip, it is obvious that the clusters of consistent statistics in the image are grouped together within the same strip.

The boundaries defined in this example (figure 3) are created by setting the MAFI input parameters such that it would segment with a measure of similarity, ζ, using dynamic programming (DP) [2]. This parameter defines the tolerance to distortion permitted during the encoding process. It is defined in equation 1:

$$\zeta = \max_i \left(\frac{\varepsilon_i^2}{\sigma_i^2 + \varepsilon_i^2} \right) \tag{1}$$

Where ε_i is the error for vector i of the segment, and σ_i^2 is the variance of the same vector.

The value of ζ should differ depending on the degree of similarity between features within each cluster, and the level of grouping required.

Figure 2: Image of a clown used as input to the coder.

Figure 3: Image of a clown with vertical lines to show boundaries set by the auto-segmentation algorithm.

Non-linear coding

Assume that there are lines of flux in an image connecting pixels of similar value. Along these lines, many pixels could be interpolated with high accuracy, and may, therefore, be omitted for coding purposes. Where there is abrupt change in texture within the image, the interpolation algorithm may reinitialise and adapt to cope with this change. MAFI is a newly developed technique, which uses DP to define these notional lines of flux. It then interpolates those lines and adapts to cope with changes within the data.

MAFI is used in coding each segmented strip of the image after it is rotated by 90°. The rotation is required because MAFI is applied orthogonally to the segmentation of the data. This is to remove the redundant, shortened, vectors, which are vertical after rotation within each strip (or horizontally, across the width of the strips before rotation).

This technique is non-linear because only those vectors that are redundant, due to their predictable characteristics, are removed from each strip . The redundancy is higher where the feature density is lower. This occurs in clustered parts of the image data. It can be eliminated by finding the cluster boundaries in one dimension (hence, the reason for auto-segmentation of the data) then removing the selected regions in the other (hence, rotation of the strips and application of MAFI for removal of redundant data). See figure 1.

Discussion

Although the selection of segmented boundaries is automatic, there is still a free choice of some parameters which will affect the sensitivity of the algorithm to changes in image statistics. This will change the number and exact location of the chosen boundaries. However, these parameters have been observed experimentally to have little effect on the overall quality of the encoded images or on the compression ratios achieved.

The quality of the image is most dependent on the maximum allowable distortion specified for the MAFI encoding of the strips. In practice, well over half the data can be labelled as redundant, and compression ratios of 50% can be achieved with little degradation in the perceived image quality.

Compare the reconstructed image of a clown in figure 4, which had been compressed by 50%, with respect to the original image in figure 2. Each strip is compressed as much as possible given the permitted level of distortion. Some strips can have much higher compression ratios, but regions such as the right-hand edge of the image reproduced here, are difficult to compress at all because of the preponderance of fine detail.

In applications such as portable video-phones or pocket televisions, where image quality is less critical, greater compression can be achieved, and ratios greater than 70% can yield quite acceptable quality. The clown in figure 5 is still recognisable after this level of compression.

It should be noted that this compression is achieved without any form of transform coding or quantisation, and the data can be further compressed by such algorithms (such as JPEG [3]).

The boundaries of the strips for both of the figures 4 and 5 are those shown in the figure 3.

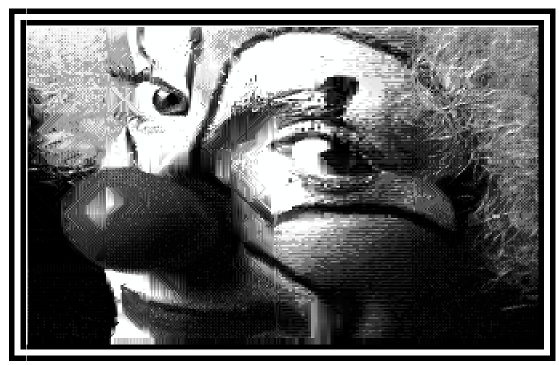

Figure 4: The clown of figure 2 encoded using the strips shown in figure 3, with a low distortion tolerance: compression ≈ 50%.

Figure 5: The clown of figure 2 encoded using the strips shown in figure 3, with a large distortion tolerance: compression ≈ 70%.

Conclusions

The advantages of encoding image data in the form of segments have been known for some time (e.g. [4, 5]). The work reported here has demonstrated that even a simple segmentation into strips can yield useful compression.

This approach provides the algorithm used to encode those strips with a more uniform distribution of image features in the transverse dimension, and they are mostly in a form which is amenable to coding by algorithms such as MAFI which treat the image as a sequence of vectors.

References

[1] BAGHAI-RAVARY, L., BEET, S. W., and TOKHI, M. O.: 'Image Coding by Multi-Step, Adaptive Flux Interpolation'. Research Report No. 592, Department of Automatic Control and Systems Engineering, The University of Sheffield, UK, 1995.

[2] NEMHAUSER, G. L.: 'Introduction to Dynamic Programming', Wiley, new York, 1966.

[3] ISO/TC97/SG2/WG8-ADCTG (1988): 'ISO Adaptive Discrete Cosine Transform Coding scheme for still image telecommunication services' (ISO adct-88, N640, January 1988)

[4] MATSUDA, I., ITOH, S. and UTSUNOMIYA, T.: 'Adaptive transform image coding based on variable-shape-blocks and directional autocorrelation models'. Signal Processing VII: Theories and Applications, September 1994, pp. 163-166.

[5] CICCONI, P., LEONARDI, R. and KUNT, M.: 'Symmetry-based image coding'. Proc. VCIP-92 Conf., Boston, SPIE-1818(3), November 1992, pp. 1312-1323.