# IMAGE-DEPENDENT SPATIAL SHAPE-ERROR CONCEALMENT

*Ferdous A. Sohel* [a], *Gour C. Karmakar* [b] and *Laurence S. Dooley* [c]

[a] School of Computer Science & Software Engineering, University of Western Australia, WA–6009, Australia
[b] Gippsland School of Information Technology, Monash University, Victoria – 3842, Australia
[c] Department of Communication and Systems, The Open University, Milton Keynes, MK7 6AA, United Kingdom
Emails: Ferdous.Sohel@csse.uwa.edu.au, Gour.Karmakar@infotech.monash.edu.au, L.S.Dooley@open.ac.uk

ABSTRACT

Existing spatial shape-error concealment techniques are broadly based upon either parametric curves that exploit geometric information concerning a shape's contour or object shape statistics using a combination of Markov random fields and *maximum a posteriori* estimation. Both categories are to some extent, able to mask errors caused by information loss, provided the shape is considered independently of the image/video. They palpably however, do not afford the best solution in applications where shape is used as metadata to describe image and video content. This paper presents a novel *image-dependent spatial shape-error concealment* (ISEC) algorithm that uses both image and shape information by employing the established rubber-band contour detecting function, with the novel enhancement of automatically determining the optimal width of the band to achieve superior error concealment. Experimental results qualitatively and numerically corroborate the enhanced performance of the new ISEC strategy compared with established shape-based concealment techniques.

## 1. INTRODUCTION

Many communication channels suffer from high error probabilities, with their deleterious result being compounded by the fact that transmitted bit-streams are normally highly compressed. The corollary of these errors is further exacerbated by the popular use of predictive and variable-length coding which can lead to error propagation. Error-resilience techniques are characterised by their innate ability to tolerate errors introduced into the bit-stream due to data loss by maintaining an acceptable picture quality, with such techniques being applied at the encoder and/or decoder. Error concealment is an important and widely used post-processing strategy, with the decoder playing the main error masking role by attempting to generate a perceptually acceptable approximation of the original data using correctly received available data.

Error concealment techniques can employ both temporal and spatial information for masking purposes, with the former for instance, exploiting the latent inter-frame correlations in video sequences to hide damaged (lost) pels in a frame by using data from already correctly received and previously concealed pels in the reference frame [1-3]. These techniques benefit from having access to past information and therefore can perform very well in a video sequence where an object's shape changes very little between consecutive frames. In contrast, when shape information changes significantly, including the emergence of new objects and object occlusion between frames, temporal methods alone are inadequate [4] and also obviously they cannot be applied to still images. In these circumstances spatial techniques are preferable, where the inherent strong neighbouring inter-pixel *spatial correlation* is exploited to conceal erroneous pels using information from correctly received and previously concealed pels within a frame [5].

978-1-4244-2179-4/08/$25.00 ©2008 IEEE

Among existing spatial shape-error concealment techniques, the *maximum a posteriori* (MAP) estimator in combination with a Markov random field has been designed for binary shape representations [6]. This exploits spatial redundant information on a statistical basis, though it does not use shape as a salient feature and so has subsequently been outperformed by those approaches which incorporate shape characteristics [5]. These typically employ parametric curves to geometrically conceal a lost boundary from correctly decoded shape information, with Bezier curves [3, 4] and Hermite splines [5] used to conceal errors, though neither has information available upon the lost parts of the contour and so depends upon just decoded shape data. Since the control points for these curves are calculated from the tangents of the two contour-ends associated with each lost segment, their performance is highly dependent on the respective tangent vectors, and since the available contours and tangents in particular, may not be representative of the lost contour, this can lead to ineffectual concealment.

Existing techniques mask spatial shape-errors independently of image information. While this may be effective in MPEG-4 based applications where shape is treated separately from texture and motion, they certainly do not provide the best solutions for a wide range of applications in which the shape is used as metadata to describe image/video content. This is because the high latent spatial correlation between the shape and its underlying image is never exploited. *Image-dependent shape coding and representation* [7], content-based image retrieval [8] and sketch-based queries using an image map to define hyperlinked objects in hyperlinked TV [9] for instance, are examples of possible applications. In these cases, shape is dependent upon the underlying image/video content and hence, the image is usually either transmitted together with the shape or accessed from an available image database at the decoder [7]. In both circumstances the received shape information may be corrupted due to data loss, so image information can be fully used for error concealment purposes. This paper presents a new *image-dependent spatial shape-error concealment* (ISEC) algorithm that utilises the image intensity gradient alongside relevant shape information. To exploit image information, ISEC employs the efficient rubber-band function [10] which detects the contour from the underlying image gradient data. Whenever a corrupted shape is received at the decoder, ISEC automatically determines the optimal width of the band, so a *region of interest* (RoI) corresponding to the lost portion of the shape is created prior to applying the rubber-band function to approximate the missing contour from the gradient data in the RoI. The performance of ISEC has been rigorously tested upon a number of popular test sequences with experimental results confirming its superior error concealment performance compared with existing techniques.

The rest of the paper is organised as follows: Section 2 presents the theoretical model for the ISEC technique including appropriate examples, while Section 3 analyses its experimental performance. Some concluding remarks are finally provided in Section 4.

## 2. IMAGE-DEPENDENT SPATIAL SHAPE-ERROR CONCEALMENT (ISEC) ALGORITHM

The two flowcharts in Fig. 1 illustrate the principal difference between traditional image-independent (Fig. 1(a)) and image-dependent shape error concealment techniques (Fig. 1(b)). While the former only exploits available shape information, the new ISEC technique utilises both shape and underlying image information.

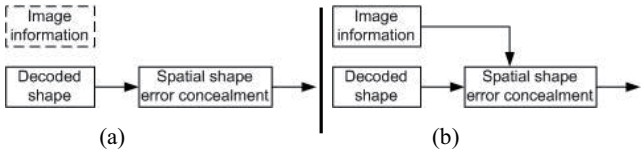



Fig. 1: Shape error concealment approaches: (a) traditional image-independent and (b) Image-dependent.

The ISEC technique, which is detailed in Fig. 2, takes both the corrupted shape that has lost blocks and the image information as its input, and generates a concealed shape representation. The process comprises a number of constituent modules namely: *contour extraction* – to extract the shape-contour from available data; *contour coupling* – to determine the associated contour end-points for each lost portion; *contour recovery* – to conceal the shape error by taking image information as an input; *shape filling* – to fill the recovered shape as though it were an original alpha plane. Each of these modules is now individually discussed.

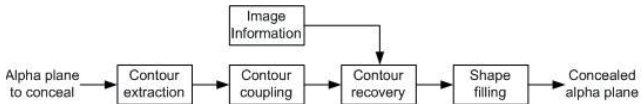



Fig. 2: The proposed ISEC technique.

2.1. ***Contour Extraction***: This occurs prior to any concealment processing and involves obtaining the correctly decoded contour from the available alpha plane. There are many different contour representations– edge, vertex and shape elements, with an edge based strategy being adopted in this paper due to its inherent robustness [4]. In this representation, the contour is considered to pass between adjacent pixels in a 4-connected neighbourhood with different values. Alternatively, the boundary of an alpha plane is a series of points that belong to the background which has at least one 4-connected neighbour belonging to the object [5].

Fig. 3 presents an illustrative example of data loss and the corresponding extracted contour. Fig. 3(a) and (b) respectively present the 1st frame of the gray scale *Bream* sequence and the corresponding binary alpha plane for the *Bream* object. Fig. 3(c) shows the alpha plane with data loss (shaded boxes) while the corresponding extracted contour is displayed in Fig. 3(d). Clearly the data loss leads to a discontinuous object contour that can be visualised as a set of independent segments.

2.2. ***Contour coupling***: Following contour extraction from the corrupted (decoded) alpha planes, the next step is to determine which of these contours actually define the missing area and the corresponding pair of available contour end-points. This is known as contour coupling, and as demonstrated in Fig 3(d), the challenge is to establish the two appropriate consecutive contour parts for each contour gap, where it is generally assumed the contour is closed and does not intersect itself [4, 5].

A contour coupling strategy based upon the total length of the concealed contour is proposed in this paper, because of its simplicity under the aforementioned assumption of no intersection in the closed contour. It needs to be highlighted however, that other contour coupling techniques like [4], are equally applicable for this purpose. The overall cost function $F$ is then defined as:

$$F = \sum_{i \in P} \left| v_1 v_2 \right|_i \tag{1}$$

where $\left|v_1 v_2\right|$ is the Euclidean distance between the candidate contour endings $v_1$ and $v_2$. $P$ is a set containing all coupled contour pairs for each coupling arrangement and $i$ represents one coupled pair in the context of the current coupling arrangement. The contour coupling that globally minimises $F$ is the final arrangement, with the respective lost data blocks also being recorded with the associated contours for each missing portion of the contour.

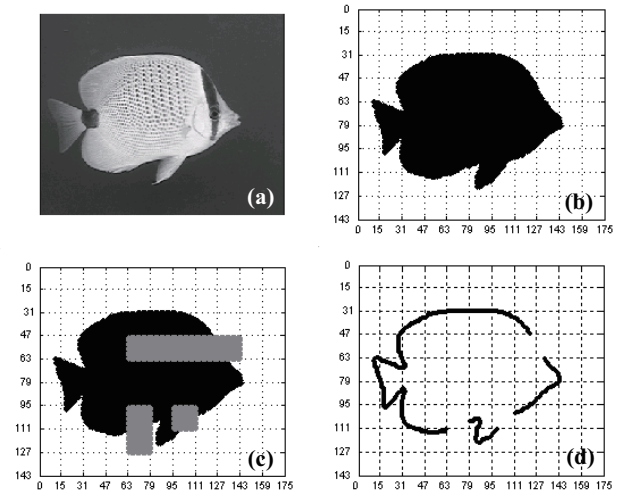



Fig. 3: (a) 1st frame of the *Bream* sequence. (b) Corresponding alpha plane of the *Bream*. (c) Received alpha plane with missing data. (d) Detected boundary for received alpha plane.

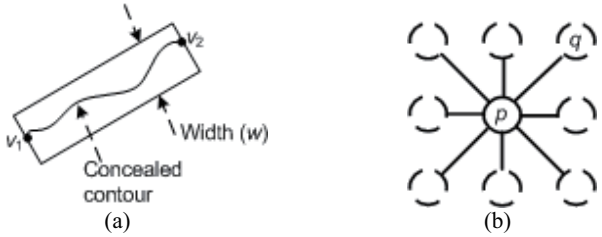



Fig. 4: (a) Illustration of the rubber-band function. (b) Edge definition of the pixel based graph.

2.3. ***Contour recovery***: The rubber-band function [10] is applied in this module. Originally proposed for image segmentation, it has subsequently been used as both an object selection tool [11] and for shape coding [7] purposes. It is a boundary detection function that performs a greedy search to maximise the total image gradient along the resulting contour. It is thus suitable for shape error concealment since the image intensity gradient is high along an object's contour [7]. To locate a contour that links the contour end points $v_1$ and $v_2$, the rubber-band function is defined as:

$$\left(v_1, v_2, w, s\right) \tag{2}$$

where $w$ is width of the rubber-band, which together with $v_1$ and $v_2$ defines the RoI as illustrated in Fig. 4(a); and $s$ is a scaling factor applied in the image gradient calculations.

The rubber-band function uses a graph search algorithm to detect boundaries. The underlying image in the RoI defined by the first three parameters of (2) is considered as a graph, where each pixel is a vertex and an 8-connected neighbourhood is considered for the edges (Fig. 4(b)). The function comprises two basic steps: (*i*) local feature computation and (*ii*) a graph search. In the first step, local image intensity gradients are calculated with a scalable edge detector operator. This gradient information is then used to define the cost function for having a weighted graph so for each graph edge $e(p,q)$, where $p$ and $q$ are neighbouring pixels, the weight (cost) function $c(p,q)$ is given by:

$$c(p,q) = \frac{1}{\|\nabla_s(q)\| + \delta} \tag{3}$$

where $\nabla_s(q)$ is the gradient at pixel $q$ with kernel size $s$ with $\delta$ being a small positive constant.

Given the cost function (3), the image is mapped into a weighted and directed graph so the overall rubber-band function in (2) becomes a shortest-path search algorithm from the source $v_1$ to the destination vertex $v_2$ which can be solved using a suitable graph searching technique, such as the Dijkstra's algorithm [12].

In general, the entire image region of the corresponding lost shape blocks associated with a coupled contour can be used as the RoI in the rubber-band function, though this is computationally inefficient since many blocks will be either in the background or lie inside the shape and so will not be part of the contour yet still incur a computational overhead in the graph search process. In these parts, the pixel-wise image-gradients are usually either very small or zero and so from (3) will incur a higher cost. To improve the computational efficiency it is necessary to determine a suitable RoI size $w$ in (2), with a Lagrangian multiplier-based optimisation approach adopted to find the optimal rubber-band width.

***Lagrangian Multiplier Optimisation:*** As $w$ increases so does the RoI, thus without loss of generality, the overall cost along contour $C(v_1, v_2)$ is a non-increasing function of $w$ whose optimal value can be obtained using a Lagrangian optimisation technique. If $I(w)$ and $C'(w)$ respectively represent an identity function and $C(v_1, v_2)$ for an arbitrary $w$, for any $\lambda \geq 0$, an unconstrained problem for the optimal solution $w^*(\lambda)$ using the *generalised Lagrangian multiplier* can be formulated as:

$$\min\left(C'(w) + \lambda \times I(w)\right) \tag{4}$$

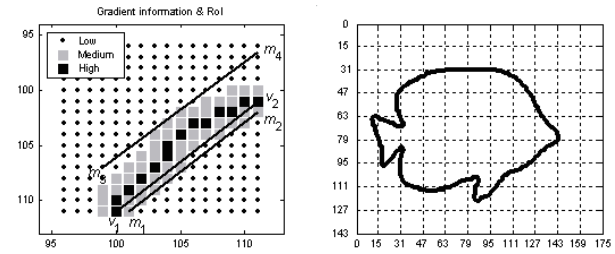
According to *Lagrangian* theory [13], the optimal solution to this unconstrained problem will also be the optimal solution to the constrained problem. So:

$$\min C'(w) \text{ subject to: } 1 \leq I(w) \leq W \tag{5}$$

where $W$ is the farthest Euclidean distance of a pixel in the coupled lost blocks from the line $v_1 v_2$. ■

While this Lagrangian approach provides the optimal $w$, it requires an iterative search to find $\lambda$. A more computationally efficient method is proposed to automatically approximate $w$ based on the image gradient, which will be shown to have negligible impact on concealment performance.

To determine $w$, the image gradient of all pixels in the corresponding image region is calculated. Gradient values are usually highest across the shape contour and then gradually decrease as the pixels move away for the contour within the lost region. Since a lower gradient value incurs a higher cost, regions having all the pixels with gradient values less than a prescribed threshold i.e. zero, can be discarded.




(a) (b)

Fig. 5: (a) Illustration of width approximation. (b) Recovered *Bream* contour.

This process is conceptualised by the example in Fig. 5(a), which shows $\nabla_4(q)$ of the lost macroblock with coordinates $((96,96),(111,111))$ in Fig. 3(a). The black squares denote pixels having *high* gradients; the gray squares *medium* gradients and dots *low* gradients. Line $v_1 v_2$ is drawn and the furthest (Euclidean distance) pixels from this line on both sides are determined. The straight line pair $m_1 m_2$ and $m_3 m_4$ which are parallel to $v_1 v_2$, pass through these furthest points as shown in Fig. 5. $m_1 m_2$ and $m_3 m_4$ together with points $v_1$ and $v_2$ then define the RoI within the corresponding image part so only those pixels inside the RoI uphold (3). It should be noted that when the gradients at the far ends of the lost block from the line $v_1 v_2$ have a high value, in the worst case the RoI becomes the entire image part corresponding to the lost blocks that are associated to the coupled contours.

The final parameter $s$ in (2) determines the scale on which the image gradient in (3) is computed. From an edge detector perspective, in general larger scales assist in attenuating texture noise while smaller scales offer better spatial localisation. In this paper $s$ is selected from the set of four values $\{1,2,4,6\}$ used in [7] in order to limit the search space for the best solution.

2.4. ***Shape filling***: After the lost contour parts have been recovered a new closed contour is obtained. The corresponding shape is then filled with shape level information, i.e., gray-scale values of 255, in what is essentially a straightforward shape filling procedure. Various approaches to achieve this goal are discussed in [14] with for simplicity, this paper using the shape filling in a non-interlaced raster scan order.

As an example, Fig. 5(b) shows the ensuing contour following the image-dependent recovery step, which after shape filling, looks exactly like the alpha plane version displayed in Fig. 3(b).

## 3. RESULTS AND ANALYSIS

There are a myriad of different ways to quantify the performance of error concealment techniques. In this paper, the popular *absolute error* (AE) and *relative error* (RE) metrics are used. The former is defined as the total number of incorrectly concealed pixels in the alpha plane, while the latter is the ratio of the incorrectly concealed pixels to the total number of lost pixels in the alpha plane [5]. As in other popular shape concealment techniques [4, 5], including MPEG-4 strategies it is assumed whenever data loss occurs, the entire macroblock ($16 \times 16$ *pel* ) is lost. Moreover, both approaches

for calculating $w$ produced almost identical concealment results, those for the approximated technique are presented in this section.

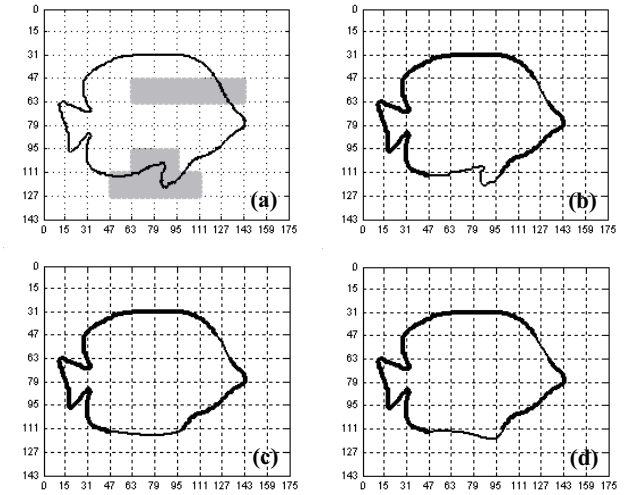



Fig. 6: (a) Shape contour of the 1st frame of the *Bream* sequence, the shaded blocks indicate the missing data. Concealment results using – (b) ISEC, (c) Bezier curve based technique [4], (d) Hermite spline based technique [5].

Fig. 6 presents the subjective performance of the new ISEC technique in comparison with two contemporary spatial shape error concealment techniques. Fig. 6(a) displays the contour of the 1st frame of the *Bream* sequence with the shaded boxes indicating lost data blocks. Fig. 6(b), (c) and (d) show the concealed contour respectively produced by the new ISEC, Bezier-curve based [4] and Hermite spline based techniques [5]. It is readily apparent from a visual comparison of all the concealed contours that ISEC provides superior performance particularly in the region of the lower fin of the object shape. The reason for this improvement is that the techniques in [4, 5] only use the geometric shape information of the coupled contours in a particular lost area, so their performance is highly dependent on the tangents therein. In contrast, ISEC employs both the underlying image as well as available shape information in its concealment strategy.

Table 1: Numerical results for error concealment applied to both *Bream* and *Akiyo* sequences for various error rates.

| Test Seq. | Error rate | *ISEC* | | *Bezier curve Technique [4]* | | *Hermite spline Technique [5]* | |
|---|---|---|---|---|---|---|---|
| | | AE (*pel*) | RE (%) | AE (*pel*) | RE (%) | AE (*pel*) | RE (%) |
| | 5% | 06 | 1.72 | 24 | 6.86 | 23 | 6.57 |
| *Bream* | 10% | 15 | 2.15 | 53 | 7.57 | 52 | 7.43 |
| (QCIF) | 20% | 48 | 3.43 | 160 | 11.43 | 162 | 11.57 |
| | 40% | 131 | 9.36 | 405 | 28.93 | 400 | 28.57 |
| | 5% | 04 | 0.84 | 12 | 2.50 | 14 | 2.92 |
| *Akiyo* | 10% | 12 | 1.25 | 29 | 3.00 | 31 | 3.20 |
| (QCIF) | 20% | 38 | 1.98 | 91 | 4.75 | 92 | 4.80 |
| | 40% | 101 | 5.26 | 225 | 11.75 | 227 | 11.84 |

Table 1 summarises the corresponding numerical results produced by the different concealment techniques upon the *Bream* and popular *Akiyo* test QCIF video sequences for various different error rates. For the *Bream* sequence, the results conclusive show an improved performance with for instance at a 10% error rate; ISEC, Bezier [4] and Hermite [5] techniques respectively generating absolute errors of 15, 53 and 52 *pels,* while the corresponding RE values are 2.15%, 7.57% and 7.43%. A similar trend is observed in the *Akiyo* results, with ISEC consistently providing comparatively superior spatial shape-error concealment performance, even at high error rates, to confirm the robustness of the new algorithm and vindicating the rationale of exploiting the underlying image information in lost regions for error concealment purposes.

## 4. CONCLUSION

This paper has presented a novel *image-dependent spatial shape-error concealment* (ISEC) technique that utilises both the shape and underlying image information for error concealment. ISEC is specifically designed for applications where an object's shape is used as dependent metadata for image descriptive purposes. It conceals shape-errors by exploiting highly correlated image information, with the kernel element being the rubber-band function which uses image data to approximate a missing contour, with the optimal width of the band being automatically determined from the image gradient information. Experimental results confirm the superior error-concealment performance of the ISEC approach compared with other existing shape-based techniques.

## References


[1] L. D. Soares and F. Pereira, "Temporal shape error concealment by global motion compensation with local refinement," *IEEE Trans. Image Proc.*, **15**(6), pp.1331-1348, 2006.

[2] G. M. Schuster and A. K. Katsaggelos, "Motion compensated shape error concealment," *IEEE Trans. Image Proc.*, **15**(2), pp. 501-510, 2006.

[3] M.-J. Chen, C.-C. Cho, and M.-C. Chi, "Spatial and temporal error concealment algorithms of shape information for MPEG-4 video," *IEEE Trans. Circuits Syst. Video Tech.*, **15**(6), pp. 778-783, 2005.

[4] L. D. Soares and F. Pereira, "Spatial shape error concealment for object-based image and video coding," *IEEE Trans. Image Proc.*, **13**(4), pp.586-599, 2004.

[5] G. M. Schuster, X. Li, and A. K. Katsaggelos, "Shape error concealment using Hermite splines," *IEEE Trans. Image Proc.*, **13**(6), pp. 808-820, 2004.

[6] S. Shirani, B. Erol, and F. Kossentini, "A concealment method for shape information in MPEG-4 coded video sequences," *IEEE Trans. Multimedia*, **2**(3), pp. 185-190, 2000.

[7] H. Luo, "Image-dependent shape coding and representation," *IEEE Trans. Circuits Syst. Video Tech.*, **15**(3), pp. 345-354, 2005.

[8] S. Berretti, A. D. Bimbo, and P. Pala, "Retrieval by shape similarity with perceptual distance and effective indexing," *IEEE Trans. Multimedia*, **2**(4), pp. 225-239, 2000.

[9] V. M. Bove Jr., J. Dakss, E. Chalom, and S. Agamanolis, "Hyperlinked TV research at MIT media laboratory," *IBM Systems Journal*, **34**(3-4), pp. 470-478, 2000.

[10] H. Luo and A. Eleftheriadis, "Designing an interactive tool for video object segmentation and annotation," in ACM International Conference on Multimedia (MM), pp. 265-269, Orlando, Florida, USA, 1999.

[11] H. Luo and A. Eleftheriadis, "Rubberband: an improved graph search algorithm for interactive object segmentation," in International Conference on Image Processing (ICIP), pp. 101-104, Rochester, New York, USA, 2002.

[12] T. H. Cormen, C. H. Leiserson, R. L. Rivest, and C. Stein, *Introduction to algorithms*, 2nd ed: The MIT Press, 2001.

[13] H. Everett, "Generalized Lagrange multiplier method for solving problems of optimum allocation of resources," *Operational Research*, vol. 11, pp. 399-417, 1963.

[14] J. D. Foley, A. van Dam, S. K. Feiner, and J. F. Hughes, *Computer graphics: principles and practice in C*, 2nd ed: Addison-Wesley Publishing Company, 1996.