SPATIAL DE-INTERLACING USING DYNAMIC TIME WARPING

Assaf Almog, Avi Levi

Technology Department,
Oplus Technologies ltd, Israel.
assaf@oplus.com ,avil@oplus.com

Alfred M. Bruckstein

Computer Science Department,
Technion – Israel Institute of Technology.
freddy@cs.technion.ac.il

ABSTRACT

Spatial de-interlacing is an essential part of motion adaptive de-interlacing used for reconstructing missing lines in cases of fast motion detection. Common spatial de-interlacing algorithms often produce artifacts in the output image, especially along edges with flat horizontal angels. In this paper we introduce a new method for spatial de-interlacing based on the Dynamic Time Warping (DTW) procedure. The DTW algorithm finds an alignment between two original consecutive lines, and then, the missing line between them is reconstructed based on this alignment. This method preserves the smoothness of the original image edges and produces a high quality progressive image.

1. INTRODUCTION

Interlaced video is used by all analogue TV broadcast systems in current use (mainly NTSC, PAL and SECAM). It is a tradeoff between the video signal's bandwidth requirements and an optimal frame rate that would not be noticeably slow for the human eye.

De-interlacing is the process of converting the interlaced video signal to a non-interlaced form in order to display it on a progressive display. Many de-interlacing techniques have been proposed (see an overview at: [1]). The most common de-interlacing algorithms are motion adaptive, i.e. inter-field interpolation is used in static scenes and intra-field interpolation is used when motion is detected. The spatial interpolation method introduced in this paper, may be used as the intra filed interpolation component of every motion adaptive de-interlacing algorithm.

Spatial de-interlacing algorithms are intra-field algorithms that interpolate the missing lines in the odd or the even fields. The missing line information is interpolated only from the available lines in the current field. These algorithms exploit the spatial correlation between the samples in the available lines and the missing lines. Naïve algorithms such as line doubling, bi-linear and bi-cubic interpolations can be easily implemented in hardware but produce poor results that can be noticed especially near diagonal edges. More elaborate algorithms use edge directed techniques ([2],[3],[4]) to produce sharper images without creating additional artifacts.

The new approach to the spatial de-interlacing problem presented in this paper is based on the Dynamic Time Warping (DTW) procedure. DTW ([5], [6]) is a method for finding the similarity between two sequences by warping the time axis of the two series in a way that achieves an optimal alignment. DTW has been successfully used for data mining, pattern recognition and speech processing. We found that DTW can be used for finding the correct alignment between two consecutive available lines. The correct alignment is determined by the image edges and therefore performing the DTW procedure yields the edge directions in those lines. Thus, the missing line can be interpolated from its two adjacent original lines according to these edge directions.

2. DYNAMIC TIME WARPING

Dynamic Time Warping is used to compute an alignment with minimal "distance" between two time series – in our case lines of pixels from a given field. The distance function is application specific. It is often the case that the two series have approximately the same overall shape, but do not line up along the time-axis. In order to find the similarity and to create the mapping between the two series, we must "warp" the time axis of one or both series so that corresponding samples appear at the same location on a common time axis. DTW is a technique for efficiently computing the optimal alignment between the two series. The alignment is optimal in the sense that it minimizes a cumulative distance measure between the series ([5] , [6]). DTW algorithms are used in many fields such as speech processing [7] and medicine [8]. In these applications it is used mainly as a method for comparing and matching two sets of information. We could not find any reference for using DTW algorithm for data interpolation. Laborelli at [9] proposed the use of the DTW procedure for removing video line jittering, which is basically a data alignment problem. However, this application does not include data interpolation.

0-7803-9134-9/05/$20.00 ©2005 IEEE

3. PROBLEM FORMALIZATION

Let A,B be two successive lines in a given interlaced field, containing N pixels each:

$$A = \{a(i) \mid i \in 1 \div N\} \;\; ; B = \{b(j) \mid j \in 1 \div N\}$$

Let L be the missing line that resides between A and B in the corresponding frame.

$$L = \{l(t) \mid t \in 1 \div N\}$$

We assume that the pixels in the image take values in a finite alphabet (i.e. 0 to 255).

We define a distance function *d(i,j)* between location i in line A and location j in line B, that may depend on the pixels values in a certain neighborhoods of the pixels *a(i)* and *b(j)*.

The elements *d(i,j)* constitute a N-by-N matrix in which the column indices relate to the pixels in line A, and the row indices relate to line B (see Figure 1).

A warping path W, is a contiguous set of ordered pairs of indices $w_k = (i_k, j_k)$ that defines a mapping between A and B:

$W = w_1, w_2, ..., w_K \;\; ; \; N \le K < 2N - 1$

The warping path W is subjected to the following constraints:

1. Boundary conditions: $w_1 = (1,1)$, $w_K = (N,N)$.
i.e. the first and last pixels in lines A and B are aligned.
2. Continuity condition:
If $w_t = (i,j)$ then $w_{t+1} = (q,p)$ where $q - i \le 1$ and $p - j \le 1$. This condition ensures smooth time warping by restricting the allowed steps in the path to adjacent elements.
3. Monotony condition:
If $w_t = (i,j)$ then $w_{t+1} = (q,p)$ where $q - i \ge 0$ and $p - j \ge 0$. This condition ensures the monotonic progress of the path w, along the time axes.

The optimal monotonic and continuous warp between A and B is the path W that minimizes the warping cost:

$$D(W) = \sum_{k=1}^{K} d(w_k)$$

The optimal distance can be efficiently found by using a dynamic programming algorithm:

$D(w_1, ..., w_l) = D(w_1, ..., w_{l-1}) +$
$\min\{d(p, q+1), d(p+1, q), d(p+1, q+1)\}$
where $w_{l-1} = (p,q)$.

The use of *d(p,q+1), d(p+1,q)* and *d(p+1,q+1)* as the only possible previous positions realizes the continuity and monotonic conditions.

Backtracking along the minimal cost pairs starting from $w_K = (N,N)$ to $w_1 = (1,1)$ yields the warping path.

In order to reduce computational burden, we use a windowing or band constraint to the warping path:

$$d(i,j) = \infty \text{ for } |i - j| > R$$

Once the DTW path is determined, the missing line L can now be calculated by interpolating its pixels from the pixels of lines A and B, according to the following rules:

1. If for the missing pixel *l(t)* there exists an element $w_n = (i,j)$ in the DTW path W such that *(i+j)* is even and $t = \frac{i+j}{2}$, then $l(t) = \frac{a(i) + b(j)}{2}$
2. If for the missing pixel *l(t)* there exist two elements $w_n = (i,j)$ and $w_{n+1} = (i+1, j+1)$ in the DTW path W such that *(i+j)* is odd and $t = \frac{i+j+1}{2}$, then $l(t) = \frac{a(i) + a(i+1) + b(j) + b(j+1)}{4}$.

It can be shown that every pixel *l(t)* falls in to one of these rules. Figure 2 demonstrates both cases. For example: pixel c in the missing line L, is interpolated using the element $w_4 = (2,4)$ and pixel j, is interpolated using the elements $w_{14} = (10,9)$ and $w_{15} = (11,10)$.

4. THE COST FUNCTION

The cost function *d(l,m)* represents the "perceptual" distance between a pair of elements (pixels) in the given lines A and B. The standard cost function in DTW algorithm is the absolute difference between the pixel values: $d(l,m) = |a_l - b_m|$. However, we found that in order to achieve better results, the cost function should be based on a more sophisticated correspondence measure between two subsets of pixels.

Let $\overline{a} = a_1 ... a_k$ and $\overline{b} = b_1 ... b_k$ be subsets of neighboring pixels of the pixels *a(i)* and *b(j)* respectively. We will use a Markovian model for the distribution of pixel values in each subset:

$P(\overline{\alpha}) = P(\alpha_1 ... \alpha_k) = P(\alpha_1) \cdot P(\alpha_2 \mid \alpha_1) \cdot P(\alpha_3 \mid \alpha_2) .... P(\alpha_k \mid \alpha_{k-1})$

, where the symbol $\overline{\alpha}$ represents $\overline{a}$ or $\overline{b}$.

For simplicity, we assume that $P(\alpha_i)$ is uniformly distributed in the corresponding alphabet, and that the conditional probability $P(\alpha_i \mid \alpha_{i-1})$ is a function of the difference between α_i and α_{i-1} only, i.e.:

$$P(\alpha_i \mid \alpha_{i-1}) = P_C(|\alpha_i - \alpha_{i-1}|)$$

Therefore:

$$(1.1)\; P(\overline{\alpha}) = P(\alpha_1) \cdot \prod_{i=2}^{k} P_C(|\alpha_i - \alpha_{i-1}|)$$

In order to construct a correspondence measure (or, equivalently, a distance function) between the two subsets we will use the Newman-Pearson score that discriminates between the following two hypotheses:

1. The subsets $\bar{a}$ and $\bar{b}$ are both distortions of a common pixel set denoted $\bar{\gamma}$.
2. The subsets $\bar{a}$ and $\bar{b}$ were sampled independently from the Markovian model.

(1.2) $$NP(\bar{a},\bar{b}) = \frac{Max_{\bar{\gamma}}\{P(\bar{\gamma}) \cdot P_D(\bar{\gamma},\bar{a}) \cdot P_D(\bar{\gamma},\bar{b})\}}{P(\bar{a}) \cdot P(\bar{b})}$$

where $P_D(\bar{\gamma},\bar{\alpha})$ is the probability of obtaining subset $\bar{a}$ as a distortion of $\bar{\gamma}$.

We assume that this probability is symmetric in $\bar{a}$ and $\bar{\gamma}$. Computing expression(1.2) involves an enumeration over all possible values of the subset $\bar{\gamma}$ and hence its computational complexity is unacceptable. A reasonable approximation for this expression assumes that the maximum is attained at one of the two options $\bar{\gamma} = \bar{a}$ or $\bar{\gamma} = \bar{b}$. Under this assumption,(1.2) can be written as:

(1.3) $$NP(\bar{a},\bar{b}) = \frac{Max\{P(\bar{a}),P(\bar{b})\} \cdot P_D(\bar{b},\bar{a})}{P(\bar{a}) \cdot P(\bar{b})} = \frac{P_D(\bar{b},\bar{a})}{Min\{P(\bar{a}),P(\bar{b})\}}$$

We now use a common assumption, consistent with (1.1), that $P(\bar{a}), P(\bar{b})$ are exponentially distributed:

$$P(\bar{a}) = Q \cdot P(a_1) \cdot e^{-\mu \sum(|a_i - a_{i-1}|)}$$

$$P(\bar{b}) = Q \cdot P(b_1) \cdot e^{-\mu \sum(|b_i - b_{i-1}|)}$$

, where Q is a normalization factor.
The distortion probability P_D is modeled by:

$$P_D(\bar{b},\bar{a}) = \prod_{i=1}^{k} C \cdot e^{-\delta|a'_i - b'_i|} \cdot e^{-\lambda\left|\sum a_i - \sum b_i\right|}$$

where: $\alpha_i' = \alpha_i - \alpha_{i-1}$ denotes the numerical derivative.
This model assumes that the distortion is a function of the difference between the horizontal derivatives of $\bar{a}$ and $\bar{b}$, and the difference between the luminance levels of the two subsets.
Using the exponential expressions for $P(\bar{a}), P(\bar{b})$ and $P_D(\bar{\alpha},\bar{b})$, we can further simplify (1.3) by considering the log of $NP(\bar{a},\bar{b})$. In conclusion, in order to define the cost function *d(l,m)* that represents the distance between two pixel neighborhood, we will use log of the correlation score (1.3) in a negative sign. Therefore, based on this model:

$$d(l,m) = \delta \cdot \left(\sum |a'_i - b'_i|\right) + \lambda \cdot \left|\left(\sum a_i - b_i\right)\right| - \mu \cdot \underset{q=a,b}{Min}\{\sum_{i=2}^{k} |q_i - q_{i-1}|\}$$

This function serves as the cost function in the DTW process, as described in section 3.

5. EXPERIMENTAL RESULTS

We have conducted a number of experiments to test the DTW method and to compare it to other spatial de-interlacing methods. We focused on the effect of the interpolation method on the overall quality of the produced image, with emphasis on the quality of the edges.
The tests were conducted for five different image samples (figure 3). Each image was first downsampled by eliminating its odd or even lines. The image was then upscaled back to its original size by three scaling methods: bi-cubic interpolation, edge-directed interpolation and DTW. The resulted images were finally compared to the original image. We used a weighted MSE measure to calculate the difference between each upscaled image and its original. The weights were determined in such a way that gives higher weights to pixels around edges.
The next table summarizes the results:

	Bi-Cubic	Edge-directed	DTW
Flower	8.4763	7.9728	7.9071
House	4.3283	3.4143	3.4382
Fabric	4.1459	1.5543	1.0226
Lighthouse	11.0966	8.7751	8.7939
Flag	2.9424	2.6742	2.6565

The quantitative results of the MSE criterion, although not sufficiently expressing the overall perceptual quality of the image, can discriminate between the edge-directed based algorithms and the linear method. The artifacts in the images that were interpolated by the bi-cubic method are easily noticeable – the edges in the interpolated pictures are blurred and seem incomplete. The resulted images of both edge-directed methods are superior in their quality to the quality of the other method – edges are smooth and complete and there are no artifacts in other parts of the images (figure 3).
The edge-directed method that was used resembles the method described in [10]. We limited both the DTW and the edge directed methods to a maximal edge distance of 6 pixels, i.e. the DTW band window was set to 6 and in the edge-directed method, the correlation was calculated between pixels that are up to a distant of 6 pixels.
Although the MSE results are similar to both edge-directed methods, a closer look at the resulted images shows the advantage of the DTW method. Fine edges appear to be more complete and artifacts near corners or crossing edges are eliminated.

6. SUMMARY AND CONCLUSIONS

In this paper, we presented a new approach to the spatial de-interlacing problem based on the DTW paradigm. We formularized a cost function that measures the similarity between two consecutive input lines, used DTW procedure to find the optimal alignment between them and interpolated the missing line's pixels, based on this alignment. The advantages of the proposed method were demonstrated in the experimental results section, where we compared its performance to the performance of existing methods such as bi-cubic and edge directed methods. To our best knowledge, the introduction of image spatial up-scaling technique based on the DTW procedure is novel. Since this method essentially performs edge detection and edge integration, it can be utilized in other image processing tasks such as segmentation, edge directed filtering and more.

7. REFERENCES

[1] G. de Haan and E.B.Bellers, "De-interlacing: an overview" *Proceedings of the IEEE*, vol. 86, pp. 1839-1857,1998.

[2] T. Doyle and M.Looymans, "Progressive scan conversion using edge information" *Signal Processing of HDTV*, II, pp. 711-721,1990.

[3] J. Salonen and S. Kalli, "Edge adaptive interpolation for scanning rate conversion" *Signal Processing of HDTV*,IV, pp. 757-764,1993.

[4] J.J. Campbell, "Adaptive diagonal interpolation for image resolution enhancement" US-patent no. 6133957, Faroudja Labs, 2001.

[5] D. Berndt and L. Clifford, "Using dynamic time warping to find patterns in time series".AAAI-94 Workshop on Knowledge Discovery in Databases (KDD-94), Seattle, Washington,1994.

[6] E. Keogh & M. Pazzani, "Derivative dynamic time warping" In proc. First International Conference on Data mining,2001.

[7] L. Rabiner & B. Juang, *Fundamentals of speech recognition* , Prentice Hall Englewood Cliffs, N.J., 1993.

[8] E.G. Caiani, A. Porta, G. Baselli, M. Turiel, S. Muzzupappa, F. Pieruzzi, C. Crema, A. Malliani & S. Cerutti, "Warped-average template technique to track on a cycle-by-cycle basis the cardiac filling phase on left ventricular volume" *IEEE Computers in Cardiology*. Vol 25 Cat. No 98CH36292, NY, USA, 1998.

[9] L. Laborelli, "Removal of video line jitter using dynamic programming approach" IEEE ICIP 2003, Barcelona, Spain, 2003.

[10] G. de Haan and R.Lodder, "De-Interlacing of video data using motion vectors and edge information", *IEEE Conference of Consumer Elec.*, pp. 70-71, Jun. 2002.

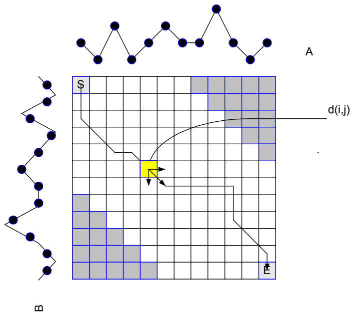

Figure 1: An example-warping path**.** Each element in the matrix represents the distance *d(i,j)* between pixel *a(i)* and *b(j).*

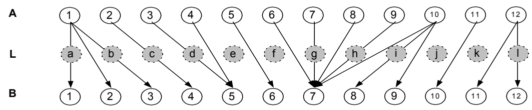

Figure 2: Missing Line Interpolation

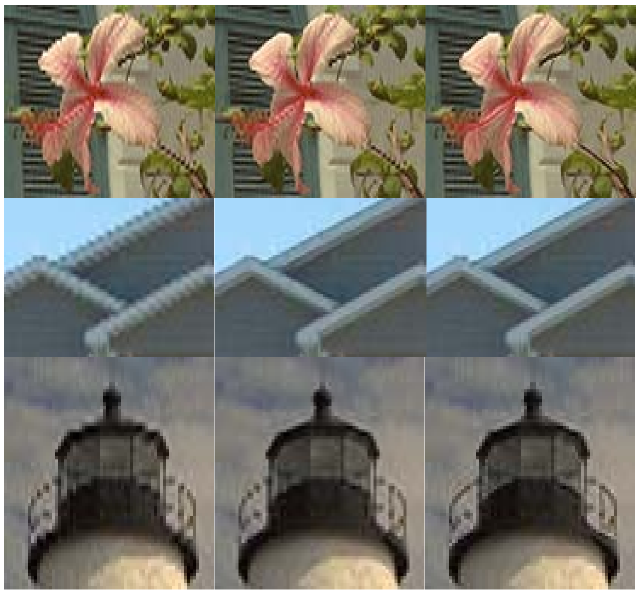

Figure 3: Resulted images- from left: bi-cubic, edge-directed and DTW. Notice the flower's stem, the house's top roof and the lighthouse's railing.