IMAGE INTERPOLATION WITH HIDDEN MARKOV MODEL

Amin Behnad, Xiaolin Wu

Department of Electrical and Computer Engineering, McMaster University
Hamilton, Ontario, Canada, L8S 4K1; behnad@grads.ece.mcmaster.ca; xwu@ece.mcmaster.ca

ABSTRACT

We propose an adaptive image interpolation technique based on hidden Markov modeling (HMM) and maximum a posterior (MAP) estimation. The HMM incorporates the statistics of high resolution (HR) images into the interpolation process and the MAP estimation exploits high-order statistical dependency between pixels. Experimental results show that the HMM-based image interpolation technique can reproduce cleaner and sharper image details than its predecessors, while suppressing common interpolation artifacts such as ringing, jaggies, and blurring.

Index Terms— Image interpolation, hidden Markov model, MAP estimation

1. INTRODUCTION

The ability of human visual system to resolve objects in a captured digital image depends on many factors, among which the spatial resolution is of great importance. High spatial resolution is necessary to reveal fine structural information on the imaged objects and scenes. High resolution directly translates to high precision in computerized image analysis, which is paramount in medical, scientific and military applications. Even in consumer electronics, entertainment industry, and other commercial applications, users desire high image resolution because resolution is in general proportional to perceived image quality.

For many applications the native resolution of a captured image is lower than required and should be enhanced. Image interpolation is the technique to compensate the inadequacy of the hardware resolution of an image. It has a wide range of applications, including resolution upconversion, resizing, video deinterlacing, video frame rate upconversion, subpixel motion estimation, image compression, etc.

The performance of an image resolution upconversion algorithm largely depends on how well it can recover high frequency components of the underlying continuous light field. Different from the task of upsampling a one-dimensional signal, upsampling an image should exploit the fact that high frequency image features, such as edges and textures, are anisotropic. A common technique adopted by many image interpolation algorithms is to estimate the edge/texture direction and interpolate in that direction [1], [2] and [3]. However, there are two issues with these algorithms. First, the edge direction is estimated from the low-resolution (LR) input image, and the estimation is error prone as insufficient sampling rate causes aliasing. Second, the edge direction is determined and directional interpolation carried out on a pixel by pixel basis disregarding the spatial coherence of edge points. In this work we overcome the above said drawbacks by applying the hidden Markov model (HMM) technique to solve the problem of image resolution upconversion. The hidden states of the HMM correspond to edge directions or smooth waveform of the underlying HR image, which are not directly observable in the LR image. Each of these HMM states is associated with an interpolator that suits the corresponding waveform. Ideally, if the HMM state was known for each pixel the corresponding interpolator would be applied for best performance. As such, image interpolation can be associated with a problem of estimating the hidden states using the observed LR image. The problem can then be solved by a maximum a posterior (MAP) sequence estimation method. This HMM-based interpolation approach has two advantages over current methods: 1) it can incorporate the statistics of HR images of a training set into the interpolation process; 2) it makes a joint decision, via MAP sequence estimation, on a block of HR pixels rather than one pixel at a time in isolation.

The rest of this paper is structured as follows. The HMM framework for image interpolation is presented in Section 2. An HMM-based image interpolator is developed in Section 3. Simulation results and comparison between the proposed method and other methods are reported in Section 4. Section 5 concludes the paper.

2. HIDDEN MARKOV MODELING FOR IMAGE INTERPOLATION

Hidden Markov models are an effective machinery of characterizing sample dependencies in Markovian processes. For image processing applications, researchers studied a two-dimensional extension of HMM (2D-HMM), and applied the 2D-HMM to problems in image classification [4] and error-resilient image communication [5]. In this research we find the 2D-HMM to well suit the task of image interpolation thanks to its ability to model two-dimensional spatial
correlations

The interpolation of a missing pixel can greatly benefit from the knowledge of the waveform of the true underlying HR image, in particular, whether the pixel is in a smooth area or near/on an edge. In the latter case it is crucial to know the edge direction. The difficulty is that the smoothness and orientation properties of the original image cannot be deduced reliably from the observed LR image. This problem has a natural HMM formulation, if we classify the original 2D intensity function at the missing pixel \( x_{i,j} \) into a set of states \( \{ \theta_0, \theta_1, \ldots, \theta_K \} \). Specifically, \( \theta_0 \) stands for the class of smooth waveforms in which an isotropic interpolator (e.g., bicubic) is effective, and \( \theta_k, 1 \leq k \leq K \), for the class of edges of direction \( k \). We want to estimate the edge directions into a small number \( K \) of classes to reduce algorithm complexity and also to avoid data overfitting in the construction of HMM using a training set. For direction \( k \), a suitable directional interpolator will be designed.

The interpolator cannot observe and thus needs to estimate the state \( s_{i,j} \in \mathbb{S} \) at pixel location \((i,j)\) that is hidden by the down sampling process. The estimation is based on observable low-resolution features exhibited by \( s_{i,j} \). In general, feature vector \( \xi_{i,j} \) of \( s_{i,j} \) consists of local attributes of the IR image in a window \( W_{i,j} \) centered at \((i,j)\).

In 2D-HMM the state \( s_{i,j} \) of a pixel \( x_{i,j} \) depends on a set of states of previous pixels [4], denoted by

\[
\Omega_{i,j} = \{ s_{t-1,j} | (i \leq t \leq i), (j \leq j \leq j) \} - \{ s_{i,j} \}.
\]  

Assuming a first-order 2D-HMM, then the conditional probability of the hidden state \( s_{i,j} \) of pixel \( x_{i,j} \) is

\[
P(s_{i,j} | \Omega_{i,j}) = \sum_{s_{i-1,j}} P(s_{i,j} | s_{i-1,j-1}, s_{i-1,j}).
\]

Using the HMM transition probabilities above, we formulate the interpolation of missing pixels on an \( M \times N \) lattice \( \mathbb{L} \) as a 2D MAP estimation problem of finding the state ensemble with maximum a posteriori probability

\[
s^* = \arg \max_s P(s | \xi) = \arg \max_s \prod_{(i,j) \in \mathbb{L}} \left\{ P(s_{i,j} | s_{i-1,j-1}, s_{i-1,j}) \right\}.
\]

The above 2D-HMM estimation framework allows off-line learning to assist image interpolation. Indeed, the 2D transition probability \( P(s_{i,j} | s_{i-1,j-1}, s_{i-1,j}) \) can be learnt from an appropriate training set of HR images whose statistics match those to be interpolated. The 2D-HMM also furnishes an adaptive image interpolator with an optimal way of exploiting 2D spatial correlations. But solving the 2D MAP problem (3) poses an algorithm challenge. Since a 2D image signal does not have a natural sequencing of the pixels, the classical dynamic programming algorithm for conventional 1D-HMM problems cannot be directly applied. One possible approach is to lump all pixels of every row (or column) into a “super-pixel” and consider the corresponding space of \( (K+1)^N \) or \( (K+1)^M \) super HMM states. Then the 2D-HMM MAP estimation problem can be converted to one of 1D MAP sequence estimation by mapping the image to a sequence of superpixels. This sequentialization would allow the use of dynamic programming algorithm. However, the above scheme is computationally intractable because the number of super states is \( O(K^{\min(M,N)}) \).

In quest for a practical algorithm for 2D-HMM based interpolation we propose to break down the 2D problem into two tightly coupled sequence estimation problems, each of which can be efficiently solved by dynamic programming. Granted such an approach can only produce an approximation solution of the 2D-HMM MAP estimation problem, but we take careful considerations not to sacrifice the use of 2D spatial correlations in our algorithm design. In estimating the 2D state ensemble we make two orthogonal scans of the image: a horizontal scan followed by a vertical scan. When scanning row \( i \), we fix the estimated states of row \( i-1 \) in (3), and compute the MAP state sequence for row \( i \)

\[
s^*_i = \arg \max_s \left\{ \sum_j \left\{ \log P(s_{i,j} | s_{i,j-1}, s_{i-1,j}) + \log p(\xi_{i,j} | s_{i,j}, s_{i,j-1}, s_{i-1,j}) \right\}, \quad s \in \mathbb{S}^N \right\}.
\]

To develop a dynamic programming algorithm to compute the MAP state sequence \( s^*_i \), we use the following recursion:

\[
w_{i,k}(1) = \log P(s_{i,1} = \theta_k) + \log p(\xi_{i,1} | s_{i,1} = \theta_k, s^*_{i-1,1}) \quad 0 \leq k \leq K
\]

\[
w_{i,k}(n) = \max_{1 \leq j < K} \left\{ w_{i,j}(n-1) + \log P(s_{i,n} = \theta_k | s_{i,n-1} = \theta_j, s^*_{i-1,n}) + \log p(\xi_{i,n} | s_{i,n} = \theta_k, s_{i,n-1} = \theta_j, s^*_{i-1,n}) \right\}
\]

\[0 \leq k \leq K, \quad 2 \leq n \leq N.\]

By solving (6) for \( n = 2, 3, \ldots, N \), we obtain

\[
\max_s P(s | \xi) = \max_{1 \leq k < K} w_{i,k}(N).
\]

and the resulting state sequence is \( s^*_i \).

The pass of row scanning produces an estimated state for each pixel \( (i,j) \), denoted by \( s_{i,j}^h \). In the next step columns are scanned and the MAP sequence estimation is performed to produce improved estimate \( s_{i,j}^b \). This is done by updating transition probabilities with the estimates of row scanning.

By now we have presented a general HMM framework for image interpolation. In the next section we will develop a new image interpolator in this framework.
3: EDGE PRESERVING IMAGE INTERPOLATION
BASED ON HMM

Now we are ready to describe a new HMM-based edge-guided image interpolator, called HMM-EGI. As illustrated by Fig. 1, the HMM-EGI algorithm reconstructs an HR image \( I_h \) from observed LR samples \( I_l \) by estimating missing pixels in two batches. In the first batch, the missing HR pixels with coordinates \((2i, 2j)\) (empty circles in Fig. 1) are interpolated. Once the missing pixels in the first batch are recovered, we obtain half of the HR pixels. In the second batch, the remaining missing pixels with coordinates \((2i, 2j - 1)\) and \((2i - 1, 2j)\) (the hatched circles in Fig. 1) are interpolated in essentially the same way as in the first batch. This is because the spatial configuration of missing pixels in the second batch becomes the same as the first batch by a 45° rotation.

Referring to Fig. 2, three hidden states are associated with each missing pixel, corresponding to two texture/edge orientations that are orthogonal to each other, plus the case of isotropic waveform (no dominant direction). For the set of missing pixels marked by empty circles in Fig. 1, positive and negative diagonal directions (see Fig. 2(a) and (b)) are the hidden states \( s_1 \) and \( s_2 \). Likewise, for the pixels marked by hatched circles, horizontal and vertical directions (shown in Fig. 2(d) and (e)) are the hidden states \( s_1 \) and \( s_2 \). In our design, cubic convolution [6] in the dominant direction is applied to interpolate missing pixels in states \( s_1 \) and \( s_2 \). The remaining hidden state \( s_0 \) is for the class of smooth waveform (see Fig. 2(c) and (f)). In this case, isotropic bicubic interpolation is applied.

To perform the MAP sequence estimation of hidden states, we seek for a feature vector that is observable in the LR image and has a high correlation to texture/edge orientation of the HR image as possible. To keep algorithm complexity low we choose a scalar feature \( \xi \) and justify our choice as follows. Due to the symmetry of the first and second batch of pixels, we define and explain the feature \( \xi \) for the first batch of pixels only in detail, and the result can be extended to the second batch straightforwardly. Recall that states \( s_1 \) and \( s_2 \) respectively represent the positive and negative diagonal directions, and hence they differ the most in waveform orientation among the three states. Given a missing pixel position \((i, j)\), let \( I_1(i, j) \) and \( I_2(i, j) \) be the results of the cubic interpolator in the positive and negative diagonal directions respectively. Since \( I_1(i, j) \) and \( I_2(i, j) \) provide two estimates of the missing HR pixel \( I_h(i, j) \), we fuse them in linear least-square sense and examine the optimal affine weight

\[
\xi = \min_x \left\{ \frac{1}{2} E\left\{ [I_h(i, j) - (x I_1(i, j) + (1-x) I_2(i, j))]^2 \right\} \right. \]

(8)

Clearly, the weight \( \xi \) approaches to 1 (or 0), as the local waveform orientation at \((i, j)\) approaches to positive 45-degree (or negative 45-degree) diagonal. If \( \xi \approx 0.5 \), then the waveform does not have a dominant orientation. Therefore, the affine weight \( \xi \in [0, 1] \) serves as a natural indicator for the three HMM states, and is defined to be the feature used by the HMM-EGI algorithm to perform MAP estimation.

However, as defined in (8), \( \xi \) is not directly observable from the LR image \( I_l \). We propose a technique to extract \( \xi \) from \( I_l \). Since \( I_h(i, j) \) is unknown, we can only compute the feature \( \xi_{i,j} \) by solving the minimization problem (8) at the locations of known pixels \( I_l(m, n) \in I_l \) in a window \( W_{i,j} \) centered at \((i, j)\). First we interpolate all missing HR pixels of the first batch in \( W_{i,j} \) in positive and negative diagonal directions separately. This generates two sets \( W_1 \) and \( W_2 \) of directional estimates for these HR pixels. Then the estimates of \( W_1 \) are used to interpolate the LR pixels \( I_l(m, n) \in W_{i,j} \) in positive diagonal direction and the results are denoted by \( I_1^2(m, n) \). Likewise, the estimates of \( W_2 \) can be used to compute \( I_2^2(m, n) \) for the negative diagonal direction. Finally, we use \( I_1^2(m, n) \) and \( I_2^2(m, n) \) to compute the feature \( \xi_{i,j} \) for the HR pixel \( I_h(i, j) \)

\[
\xi_{i,j} = \min_x \left\{ \frac{1}{2} \sum_{W_{i,j}} [I_l(m, n) - x I_1^2(m, n) - (1-x) I_2^2(m, n)]^2 \right\} .
\]

(9)

Given the hidden states of HMM the values of feature \( \xi_{i,j} \) are assumed to obey a Gaussian distribution

\[
p(\xi_{i,j} | s_{i,j}, s_{i,j-1}, s_{i-1,j}) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(\xi_{i,j} - \mu)^2}{2\sigma^2}}.
\]

(10)

For each combination of \( s_{i,j}, s_{i,j-1}, s_{i-1,j} \) (27 in total) the mean \( \mu \) and variance \( \sigma^2 \) are estimated by off-line training with a large set of HR images.

4. SIMULATION RESULTS

The proposed multidirectional image interpolation method was implemented and compared with a number of existing methods: bilinear, bicubic [1], method of [2], and
Fig. 2 Three waveform orientations corresponding to hidden states \( s_1 \), \( s_2 \), and \( s_3 \) for the first batch (top) and the second batch (bottom).

<table>
<thead>
<tr>
<th>Table 1</th>
<th>PSNR (decibels) results of reconstructed images for methods bicubic, NEDI [2] and the proposed HMM-EGI</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Image</strong></td>
<td>bicubic</td>
</tr>
<tr>
<td>Lena</td>
<td>33.58</td>
</tr>
<tr>
<td>Flowers</td>
<td>27.65</td>
</tr>
<tr>
<td>Motor</td>
<td>38.34</td>
</tr>
<tr>
<td>Window</td>
<td>37.01</td>
</tr>
<tr>
<td>Parrot</td>
<td>36.35</td>
</tr>
<tr>
<td>Average</td>
<td>34.58</td>
</tr>
</tbody>
</table>

MRF-EDI [4], classic bicubic interpolator [6] and the NEDI interpolator [2]. The PSNR results of the three methods are presented in Table 1. As shown by the table the proposed method achieves appreciably better objective performance in term of PSNR than the other five methods.

To evaluate the performance in visual quality, we present in Fig. 3 a side-by-side close up comparison between NEDI and HMM-EGI. The upconverted images by HMM-EGI apparently have sharper and cleaner edges and details.

5. CONCLUSION

A new adaptive interpolation algorithm based on HMM is proposed (HMM-EGI). In this technique, the problem of image interpolation is converted to one of MAP sequence estimation. HMM-EGI incorporates the statistics of HR images, which are supplied by a training set, into the interpolation process and is able to exploit high-order statistical dependency between missing pixels. A comparison study demonstrates the competitive performance of the HMM-EGI interpolation technique in both visual quality and PSNR metric.

6. REFERENCES


