AUTOMATIC PARAMETRISATION FOR AN IMAGE COMPLETION METHOD BASED ON MARKOV RANDOM FIELDS

Huy Tho Ho[*]

School of Electrical and Electronic Engineering
University of Adelaide
Adelaide, Australia
huy.ho@student.adelaide.edu.au

Roland Goecke

NICTA Canberra Research Laboratory[†]
RSISE, Australian National University
Canberra, Australia
roland.goecke@anu.edu.au

ABSTRACT

Recently, a new exemplar-based method for image completion, texture synthesis and image inpainting was proposed which uses a discrete global optimization strategy based on Markov Random Fields. Its main advantage lies in the use of priority belief propagation and dynamic label pruning to reduce the computational cost of standard belief propagation while producing high quality results. However, one of the drawbacks of the method is its use of a heuristically chosen parameter set. In this paper, a method for automatically determining the parameters for the belief propagation and dynamic label pruning steps is presented. The method is based on an information theoretic approach making use of the entropy of the image patches and the distribution of pairwise node potentials. A number of image completion results are shown demonstrating the effectiveness of our method.

***Index Terms*—** Image restoration, Markov processes, Stochastic fields

1. INTRODUCTION

Image completion is the process of filling the unknown region of an image from its observed part in a visually plausible way. It has been an active research topic in recent years and there have been many advances in the development of algorithms for solving this problem. Some examples are a statistical-based method [1], a PDE-based method [2] that propagates image Laplacians in the isophote direction, and an exemplar-based method [3] that synthesizes pixels or image patches using texture synthesis techniques. A recently proposed exemplar-based technique [4] considers the image completion problem as a discrete global optimization with a well defined objective function based on a Markov Random Field (MRF). Thus, this technique overcomes the limitations of other approaches such as greediness and ineffectiveness in completing images where complex structures exist in the unknown region. This approach also carries two important improvements over standard belief propagation (BP): *priority-based message scheduling* and *dynamic label pruning* to significantly reduce the performing time of BP.

This paper extends the algorithm used in [4] into a more generic framework by providing an automatic mechanism to determine parameters needed for the completion process, so that it can be performed automatically on any input image without user intervention. The remainder of this paper is organised as follows. Section 2 gives an overview of related work and describes how Markov Random Fields and Belief Propagation are applied to solve the problem of image inpainting, in particular the improvements in the algorithm in [4] and their effects in increasing the speed of BP. In Section 3, our method of automatically determining the required parameters is proposed. Results of testing input images with different parameters are presented in the Section 4. Finally, the conclusions are given in Section 5.

2. RELATED WORK

The method proposed in [4] has significantly reduced the time to perform image completion by using MRFs and BP. According to the authors, their method reduced the computation time for input images of size 320×240 pixels from hours to only a few minutes on a P4 2.4GHz CPU, compared to standard BP. On the other hand, the method is quite parameter-dependent and different sets of parameters lead to different outputs for the same input image. Moreover, the parameters need to be chosen manually in a heuristic manner which is not desirable.

2.1. Markov Random Fields

The general framework for the image completion problem will be defined as follows (using the same notations as in [4]). Let I_0 be the input image with a target region T and a source region S ($S \subseteq (I_0 - T)$). We want to fill the region T using

[*]Huy Tho Ho performed the work reported in this paper while at the Australian National University and the NICTA Canberra Research Laboratory.

[†]National ICT Australia is funded by the Australian Government's *Backing Australia's Ability* initiative, in part through the Australian Research Council.

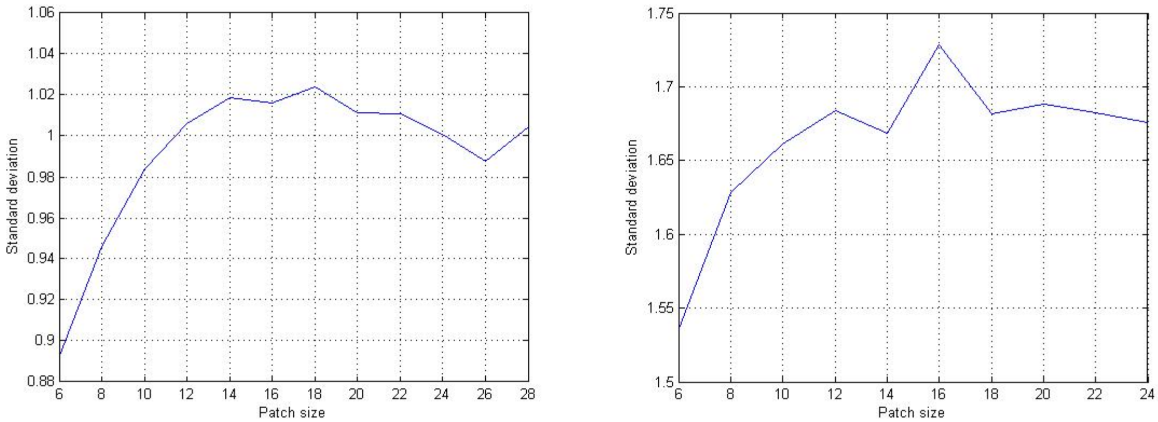
(a)  (b)

Fig. 1: Standard deviation of the patch entropies versus patch size. (a) Indoor scene. (b) Outdoor scene.

patches from S in a visually plausible way by using MRFs.

The image I_0 will be divided into a lattice with horizontal and vertical spaces (gap_x, gap_y) between nodes, respectively. A lattice point is a MRF node n if its $w \times h$ neighborhood intersects T. A 4-neighborhood system is created by edges E of the MRF nodes $\{n\}_{i=1}^{N}$. A set of labels L is formed in S containing all $w \times h$ patches l that do not intersect with T.

We define the following energy function that determines the quality of a labeling:

$$E(\{l_i\}) = \sum_{i=1}^{N} V_i(l_i) + \sum_{(i,j)\in E} V_{ij}(l_i, l_j) \qquad (1)$$

where

$$V_i(l) = \sum_{p\in[-\frac{w}{2}\,\frac{w}{2}]\times[-\frac{h}{2}\,\frac{h}{2}]} M(n_i+p)(I_0(n_i+p) - I_0(l+p))^2 \qquad (2)$$

V_i is called a *single node potential*. The pairwise potential $V_{ij}(l_i, l_j)$ is defined similarly as the cost of assigning labels (l_i, l_j) to two neighboring nodes (n_i, n_j) and is calculated as the sum of squared differences (SSD) over the overlapping region of the two labels. It is also known as compatibility function as it measures how well these patches agree [5]. The optimal labeling $\widehat{L} = \{\widehat{l_i}\}_{i=1}^{N}$ is found by minimizing Eq. (1).

2.2. Priority Belief Propagation and Label Pruning

Using negative logarithmic probabilities, a message from node n_i to a neighboring node n_j at time t is defined as:

$$m_{ij}^{t}(l) = \min_{l_i\in L}\{V_i(l_i) + V_{ij}(l_i, l_j) + \sum_{k:k\neq j,(k,i)\in E} m_{ki}^{t-1}(l_i)\} \qquad (3)$$

After s iterations, a belief vector $\{b_i(l)\}_{l\in L}$ is computed and the label $\widehat{l_i} = \arg\max_{l\in L} b_i(l)$ is selected individually at each node

$$b_i(l) = -V_i(l) - \sum_{k:(k,i)\in E} m_{ki}^{s}(l) \quad . \qquad (4)$$

However, standard BP is slow [5], heuristic [6], and requires user intervention [7]. In [4], two improvements to BP were introduced to increase the speed and to make the algorithm converge after a small number of iterations.

The first extension is the use of *message scheduling*. The order of a node transmitting messages to its neighbors will be determined based on the confidence about its labels. This will help the node that has the most informative messages transmit first in order to increase the confidence of it neighbors. As a result, its neighbors will be more tolerable to label pruning and the algorithm converges faster. The priority of a node is defined as the inverse of the cardinality of set P: $\text{priority}(n_i) = \frac{1}{|P|}$ where $P(n_i) = \left|\{l \in L : b_i^{rel}(l) \geq b_{conf}\}\right|$, b_{conf} is the confidence threshold belief, $b_i^{rel}(l) = b_i(l) - b_i^{max}$ is the relative belief, and b_i^{max} is the maximum belief of node n_i.

The second improvement is *dynamic label pruning*. Label pruning will be applied to a node if its number of active labels is greater than L_{max}, a user specified constant. The labels of a visited node are traversed in descending order of relative belief and those with $b_i^{rel}(l) \geq b_{prune}$ are chosen as active labels for this node. b_{prune} is the label pruning threshold belief. In order to avoid choosing many similar labels to the active label set, a label is declared as active only if it is not too similar to any of the already active labels, i.e. the SSD between this label and any of the chosen labels is less than a threshold $SSD_{similar}$. Note that a minimum number of labels L_{min} is always kept for each node. Applying label pruning to BP helps reducing the complexity of updating the messages from $O(|L|^2)$ to $O(|L_{max}|^2)$ [4].

3. AUTOMATIC PARAMETRISATION

In this section, a way for automatically determining the required parameters of the image completion method is presented. First, we present a method for determining the optimal patch size based on the information content in the patches.

Then, we use the distribution of node potentials to automatically determine the label pruning parameters.

3.1. Patch Size

In the process of completing an image using MRFs and BP, determining the right size of the patches helps improving the quality of the output. If the patch size is too small, the patches do not contain sufficient information for estimating the underlying scene texture. Otherwise, if the chosen patch size is too large, block effects appear in the output and it is also difficult to learn the relationship between the local node and the scene patches [5]. In this section, we propose a method of choosing the right patch size based on the distribution of the information content of the label patches.

The *entropy* of a label patch is defined as:

$$e = -\sum_{i=0}^{G-1} p(z_i) \log p(z_i) \quad (5)$$

where z is a random variable denoting gray levels, G is the number of distinct gray levels, and $p(z_i), i = 0, 1, 2, \ldots, G-1$ is the corresponding histogram of the label patch.

The *entropy* of a label patch defines the average amount of information obtained by observing that patch. It will have a small value if the dominant property of that patch is tone and large value if the dominant property is texture [8]. To create a set of labels that is the most suitable for the image completion task, the amount of information in each patch is not as important as the distribution of information across patches in the label set.

Let $\Omega_u = \{e_1, e_2, e_3, \ldots, e_N\}$ be the set of entropy values for patches of u. N is the number of labels obtained for this size. Assuming that the distribution of values across all Ω is Gaussian, the criterion for choosing the optimal patch size is equivalent to maximizing the standard deviation $arg\ max_{u,\Omega}\ \sigma(\Omega_u)$. By maximizing the statistical dispersion of the entropy values, we ensure that the best choice of patches is available for the image completion task.

3.2. Label Pruning Parameters

In the algorithm proposed by [4], parameters such as b_{conf}, b_{prune} and $SSD_{similar}$ are determined in a heuristic way. From Section 2.2, we know that the label pruning will be performed only at the first iteration for each node and from Equations (1)-(4) we know that the potentials are a good choice of representing the beliefs of each node. With a chosen patch size of $w \times h$, the SSD of each pair of labels will be calculated to get a histogram. For a pair of labels l_i and l_j, the SSD is defined as:

$$SSD(l_i, l_j) = \sum_{p \in [-\frac{w}{2}\ \frac{w}{2}] \times [-\frac{h}{2}\ \frac{h}{2}]} (I_0(l_i + p) - I_0(l_j + p))^2 \quad (6)$$

As the number of labels is large, in order to increase the speed of calculating the histogram, the SSD values will be divided into Φ intervals, i.e. bins, of width q. Let n_k be the number of SSDs that lie in bin k, i.e. $q\,k < SSD \leq q\,(k+1)$, and n is the total number of values of SSD, the probability that a SSD value is in the interval $[q\,k, q\,(k+1)]$ is:

$$Pr(k) = \frac{n_k}{n} \qquad k = 0, 1, 2, ..., \Phi - 1 \quad (7)$$

From these $Pr(k)$, the parameters b_{conf}, b_{prune}, and $SSD_{similar}$ are chosen such that:

$$\sum_{k=0}^{-b_{conf}/q} p_r(k) = \frac{1}{2} \quad (8)$$

$$\sum_{k=0}^{-b_{prune}/q} p_r(k) = \frac{3}{4} \quad (9)$$

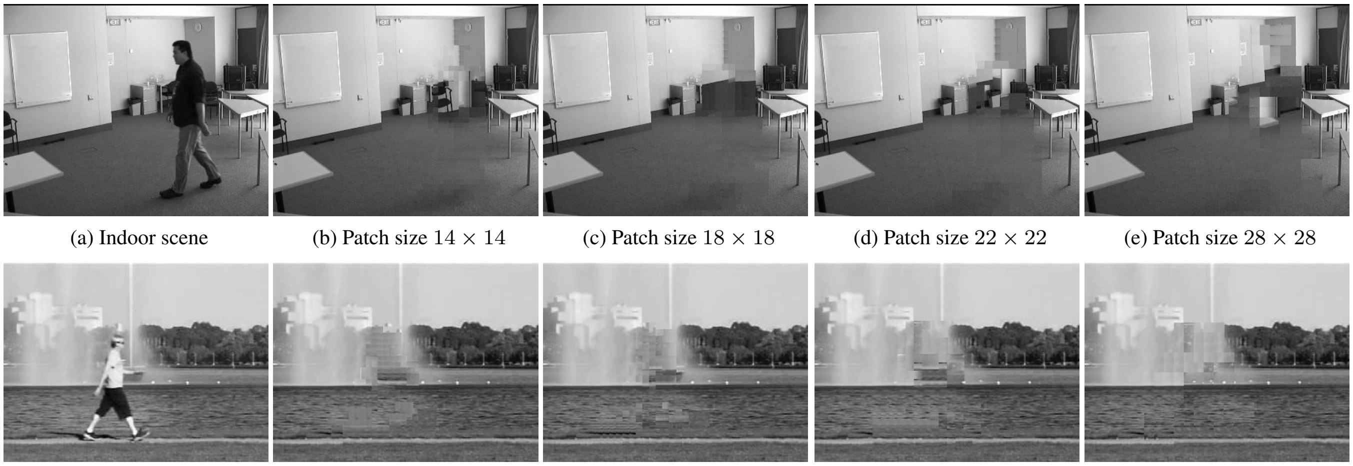
$$\sum_{k=0}^{SSD_{similar}/q} p_r(k) = \frac{1}{4} \quad (10)$$

In other words, b_{conf} is chosen so that the cumulative sum of the probabilities $Pr(k)$ from bin $k = 0$ to bin $k = \frac{-b_{conf}}{q}$ is 0.5, i.e. 50% of the total number of SSDs are smaller than $-b_{conf}$. In the same manner, 75% of the total number of SSDs will be greater than $-b_{prune}$ because we do not want to over-prune good labels. The value of $SSD_{similar}$ is chosen to be greater than 75% of the SSDs to help avoid wasting a large part of the L_{max} labels for similar patches.

4. EXPERIMENTS AND DISCUSSION

The proposed automatic parametrisation approach was tested on a variety of images taken in both indoor and outdoor conditions (see Fig. 2 for two examples). In the experiments, we chose $w = h$ and $gap_x = gap_y = \frac{1}{2}w = \frac{1}{2}h$ to get a good overlap region between two labels. As mentioned above, it is sensible to limit the possible size of patches to be between a minimum and a maximum patch size. In our experiments, the minimum patch size is 6×6 and the maximum is $\frac{1}{10}$ of the size of the image. For a patch size below the minimum, we found that patches did not contain sufficient information for the image completion method to produce perceptually good results. Similarly, a patch size larger than the maximum resulted in strong, undesirable 'block' effects. For determining the label pruning parameters, a bin size of $q = 10^2$ was used.

First, the optimal patch size was determined (cf. Sec. 3.1). Fig. 1 shows the graphs of standard deviation of patch entropy versus patch size. Once the optimal patch sizes were known, which were 18×18 for the indoor scene and 16×16 for the outdoor scene, the label pruning parameters were determined (Sec. 3.2). Table 1 and Fig. 2 show the resulting parameters and output images, respectively.



(a) Indoor scene (b) Patch size 14×14 (c) Patch size 18×18 (d) Patch size 22×22 (e) Patch size 28×28

(f) Outdoor scene (g) Patch size 14×14 (h) Patch size 16×16 (i) Patch size 20×20 (j) Patch size 24×24

Fig. 2: Completion results for different patch sizes

Patch size	b_{conf}	b_{prune}	$SSD_{similar}$
14	-223×10^2	-566×10^2	5×10^2
18	-371×10^2	-895×10^2	25×10^2
22	-533×10^2	-1241×10^2	48×10^2
28	-874×10^2	-1995×10^2	120×10^2

(a) Parameters for the indoor scene

Patch size	b_{conf}	b_{prune}	$SSD_{similar}$
14	-263×10^2	-572×10^2	10^2
16	-328×10^2	-733×10^2	6×10^2
20	-507×10^2	-1114×10^2	17×10^2
24	-704×10^2	-1577×10^2	41×10^2

(b) Parameters for the outdoor scene

Table 1: Parameters determined for different patch sizes

For the indoor scene, the floor area was filled in well in all output images. However only in Fig. 2c, the table and the wall were completed in a plausible way. Perceptually, it is clear that patch size 18 provided the best result for the indoor scene, which is in accordance with the result of the information theoretic approach (Fig. 1). For the outdoor scene (Fig. 2f), the water surface and the pathway were filled in well in all the outputs. However, the fountain area was reconstructed convincingly only for patch size 16, as predicted in Fig. 1.

5. CONCLUSIONS

An automatic parametrisation method for the image completion method proposed by [4] has been presented. Through the variation of the information content in each set of labels, the optimal size of patches can be estimated. The parameters of the Priority BP are calculated from the distribution of node potentials. Our approach provides a framework for solving the image completion task without the need to manually choose parameters. In future, we plan to use higher order MRFs for image completion of complex scenes.

6. REFERENCES

[1] J. Protilla and E.P. Simoncelli, "A parametric texture model based on joint statistics of complex wavelet coefficients," *International Journal of Computer Vision*, vol. 40, no. 1, pp. 49–70, 2000.

[2] M. Bertalmio, G.Sapiro, V. Caselles, and C. Ballester, "Image Inpainting," in *Proc. ACM SIGGRAPH 2000*, 2000, pp. 417–424.

[3] A.A. Efros and T.K. Leung, "Texture synthesis by non-parametric sampling," *Proceedings of ICCV'99*, vol. 2, pp. 1033–1038, 1999.

[4] N. Komodakis and G. Tziritas, "Image Completion Using Global Optimization," in *Proc. CVPR2006*, New York, USA, June 2006, vol. 1, pp. 442–452.

[5] W.T. Freeman, E.C. Pasztor, and O.T. Carmichael, "Learning Low-level vision," *International Journal of Computer Vision*, vol. 40, no. 1, pp. 25–47, 2000.

[6] P.F. Felzenszwalb and D.P. Huttenlocher, "Efficient Belief Propagation for Early Vision," *International Journal of Computer Vision*, vol. 70, no. 1, pp. 41–54, 2006.

[7] J. Sun, L. Yuan, J. Jia, and H.-Y. Shum, "Image completion with structure propagation," in *Proceedings of the ACM SIGGRAPH*, 2005, vol. 24, pp. 861–868.

[8] R.M. Haralick, "Statistical and structural approaches to texture," *Proceedings of the IEEE*, vol. 67, no. 5, pp. 786–804, May 1979.