

Quaternion Potential Functions for a Colour Image Completion Method Using Markov Random Fields

Huy Tho Ho
School of Electrical and Electronic Engineering
The University of Adelaide
Adelaide, Australia
huy.ho@adelaide.edu.au

Roland Goecke
RSISE
Australian National University
Canberra, Australia
roland.goecke@anu.edu.au

Abstract

An exemplar-based algorithm has been proposed recently to solve the image completion problem by using a discrete global optimisation strategy based on Markov Random Fields. We can apply this algorithm to the task of completing colour images by processing the three colour channels separately and combining the results. However, this approach does not capture the correlations across the colour layers and, thus, may miss out on information important to the completion process. In this paper, we introduce the use of quaternions or hypercomplex numbers in estimating the potential functions for the image completion algorithm. The potential functions are calculated by correlating quaternion image patches based on the recently developed concepts of quaternion Fourier transform and quaternion correlation. Experimental results are presented for image completion which evidence improvements of the proposed approach over the monochromatic model.

1. Introduction

Image completion has been a challenging and active research topic in image processing and computer vision in recent years. It is the process of replacing the unknown region of an image by textures from the observed part in a visually plausible way. Figure 1 shows an example.

There have been many advances in the development of algorithms for solving this problem. Some examples are a statistical-based method [10], a PDE-based method [1] that propagates image Laplacians in the isophote direction, and an exemplar-based method [2] that synthesises pixels or image patches using texture synthesis techniques. Another recently proposed exemplar-based technique [8] considers the image completion problem as a discrete global optimisation with a well defined objective function based on a Markov

Random Field (MRF) and uses Belief Propagation (BP). 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 BP: *priority-based message scheduling* and *dynamic label pruning* to significantly reduce the time to perform the BP.

The algorithm in [8] can be applied to completing colour images by processing each colour channel separately and combining the results. However, this approach does not exploit the correlations between the different channels and, thus, may leave out information that may be essential to obtain a visually plausible result. This paper proposes a new, systematic way of applying the algorithm of [8] to colour images by applying quaternions in calculating the required potential functions. Quaternions are hypercomplex numbers which have a real part and three orthogonal imaginary parts. The remainder of the paper is organised as follows. Section 2 gives an overview of how MRFs and BP are applied to solve the problem of image completion, in particular the improvements in the algorithm in [8] and their effects in increasing the speed of BP computation. In Section 3, we will describe some basics of quaternion numbers and their Fourier transform. Next, in Section 4, our method of applying quaternions in the potential functions is proposed. In Section 5, we show a number of experimental results to demonstrate the effectiveness of our approach. Finally, the conclusions are given in Section 6.

2. Image Completion by Global Optimisation

2.1. Markov Random Fields

By using the same notations as in [8], we can define the general framework for the image completion problem as follows. Let I_0 be the input image with a target region



(a) Original image



(b) Completion output

Figure 1. An example of image completion. The foreground person was manually segmented and then automatically removed.

T and a source region S ($S \subset I_0 - T$). Let M be the binary mask, non-zero only in region S . The target region T will need to be filled by copying patches from a label set L formed in S in a visually plausible way using a discrete MRF.

The input image I_0 will be divided into a lattice with horizontal and vertical spaces (gap_x, gap_y) between nodes, respectively. Lattice points whose $w \times h$ neighbourhood intersects the target region T will form a set of MRF nodes $\{n_i\}_{i=1}^N$. A 4-neighbourhood system is then created by edges E of the MRF. Figure 2 shows the binary mask M and the resulting discrete MRF for Figure 1a. Each label l in the label set L is a $w \times h$ patch from S that does not intersect with T .

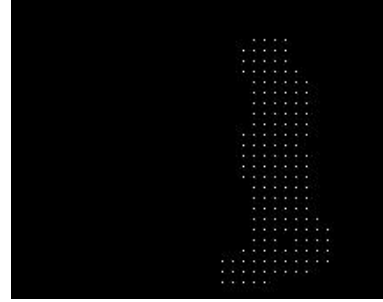
The *single node potential*, V_i , for placing a patch l over node n_i is defined as

$$V_i(l) = \sum_{p \in [-\frac{w}{2}, \frac{w}{2}] \times [-\frac{h}{2}, \frac{h}{2}]} M(n_i + p)(I_0(n_i + p) - I_0(l + p))^2 \quad (1)$$

where V_i is an indication of the similarity between the patch and the region around n_i . The pairwise potential $V_{ij}(l_i, l_j)$ is defined similarly as the cost of assigning labels (l_i, l_j) to two neighbouring nodes (n_i, n_j) and is calculated as the sum of squared differences (*SSD*) over the overlapping re-



(a) Binary mask



(b) MRF nodes

Figure 2. Binary mask and the MRF nodes for Figure 1a.

gion of the two labels. It is also known as the compatibility function as it measures how well these patches agree [5]. The optimal labelling $\hat{L} = \{\hat{l}_i\}_{i=1}^N$ is found by minimising the energy function

$$E(\{l_i\}) = \sum_{i=1}^N V_i(l_i) + \sum_{(i,j) \in E} V_{ij}(l_i, l_j) \quad (2)$$

2.2. Priority Belief Propagation and Label Pruning

BP is an optimisation technique that works by passing local messages along the nodes of a MRF [4]. A message from node n_i to a neighbouring node n_j at time t is defined using negative logarithmic probabilities

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

Assuming that all messages stabilise after s iterations (convergence), the label $\hat{l}_i = \arg \max_{l \in L} b_i(l)$ that maximises the belief is selected individually at each node. $b_i(l)$, the belief of node i for label $l \in L$, is computed as

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

However, standard BP is slow [5], heuristic [4], and requires user intervention [12]. In [8], 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 confidence of a node about its labels is used to determine the transmitting order for that node. *The node most confident about its labels should be the first one to transmit outgoing messages to its neighbours* [8]. This scheduling principle will help the node that has the most informative messages transmit first in order to increase the confidence of its neighbours. As a result, the neighbours will be more tolerable to label pruning. The algorithm also converges faster after a small fixed number of iterations. The priority of a node is defined as the inverse of the cardinality of set P $\text{priority}(n_i) = \frac{1}{|P|}$ where $P(n_i) = |\{l \in L : b_i^{\text{rel}}(l) \geq b_{\text{conf}}\}|$, b_{conf} is the confidence threshold belief, $b_i^{\text{rel}}(l) = b_i(l) - b_i^{\text{max}}$ is the relative belief, and b_i^{max} is the maximum belief of node n_i .

The second improvement in [8] is *dynamic label pruning*. This process is applied to a node if the number of active labels for that node is greater than L_{max} , a user specified constant. When a node is visited, its labels are traversed in descending order of relative belief and those with $b_i^{\text{rel}}(l) \geq b_{\text{prune}}$ are marked as active. b_{prune} is the label pruning threshold belief. Furthermore, a label is declared as active only if it is not too similar to any of the already active labels in order to avoid choosing many similar labels for a node. As a consequence, the *SSD* between this label and any of the other chosen labels must be 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 $O(|L_{\text{max}}|^2)$ which is still quadratic but as $L_{\text{max}} \ll L$, the computation time is significantly reduced [8]. The speed of BP can also be enhanced by pre-computing the reduced matrices of pairwise potentials.

3. Quaternion Numbers

3.1. Introduction to Quaternion Numbers

Quaternions or hypercomplex numbers are a non-commutative extension of complex numbers to four dimension. They were first introduced by Hamilton in [6]. Following [14], a quaternion q is a number with a scalar $S(q)$ real part and a vector $v(q)$ imaginary part. In Cartesian form, q can be written as

$$q = S(q) + V(q) = a + bi + cj + dk \quad (5)$$

where $a, b, c,$ and d are all real and \mathbf{i}, \mathbf{j} and \mathbf{k} are orthogonal imaginary operators [9] that obey

$$\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = \mathbf{ijk} = -1 \quad (6)$$

$$\mathbf{ij} = -\mathbf{ji} = \mathbf{k} \quad \mathbf{jk} = -\mathbf{kj} = \mathbf{i} \quad \mathbf{ki} = -\mathbf{ik} = \mathbf{j} \quad (7)$$

The quaternion conjugate and modulus of q are given by

$$\bar{q} = a - bi - cj - dk \quad (8)$$

$$|q| = \sqrt{a^2 + b^2 + c^2 + d^2} \quad (9)$$

If $a = 0$, q is a pure quaternion. If $|q| = 1$, it is a unit quaternion.

Given two pure quaternions \mathbf{u} and \mathbf{v} , \mathbf{u} may be resolved into components *parallel* and *perpendicular* to \mathbf{v} as below [9]

$$\mathbf{u}_{\parallel} = \frac{1}{2}(\mathbf{u} - \mathbf{vuv}), \quad \mathbf{u}_{\parallel} \parallel \mathbf{v} \quad (10)$$

$$\mathbf{u}_{\perp} = \frac{1}{2}(\mathbf{u} + \mathbf{vuv}), \quad \mathbf{u}_{\perp} \perp \mathbf{v} \quad (11)$$

The above equations can also be extended to a full quaternion q such that $q_{\parallel} = S(q) + \mathbf{V}_{\parallel}(q)$ and $q_{\perp} = \mathbf{V}_{\perp}(q)$. If q is a full quaternion and \mathbf{p} is a vector where $\mathbf{V}(q) \perp \mathbf{p}$, we can reorder them such that $q\mathbf{p} = \mathbf{p}\bar{q}$. As a result, it can be seen that parallel quaternions commute. This result is important for the practical form of quaternion correlation [14].

A colour image in RGB space may be represented using hypercomplex numbers by encoding the three colour components of the image as a pure quaternion:

$$f(x, y) = r(x, y)\mathbf{i} + g(x, y)\mathbf{j} + b(x, y)\mathbf{k} \quad (12)$$

where $r(x, y), g(x, y), b(x, y)$ are the red, green, and blue components at the coordinate (x, y) , respectively. This representation is chosen because a point in an RGB image represents a 3-space vector as does the pure quaternion [9].

3.2. Quaternion Fourier Transform

From the definition of a quaternion, the earliest Quaternion Fourier Transform (QFT) was introduced in [3]. There are many different QFT formulations available. The one defined in [9] is used in this paper as it has been proven to be the best so far for the computation of correlation which will be relevant in our approach. Because quaternion multiplication is not commutative, there are two different QFTs for a quaternion function $f(\mathbf{x})$: the left side transform $F^L[f(\mathbf{x})]$ and the right side transform $F^R[f(\mathbf{x})]$. They are defined as

$$\begin{aligned} F^{\pm L}[f(\mathbf{x})] &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{\mp \mu \mathbf{u}^T \mathbf{x}} f(\mathbf{x}) d\mathbf{x} = F^{\pm L}[\mathbf{u}] \Leftrightarrow \\ F^{\mp L}[F^{\pm L}[\mathbf{u}]] &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{\pm \mu \mathbf{u}^T \mathbf{x}} F^{\pm L}[\mathbf{u}] d\mathbf{u} = f(\mathbf{x}) \\ F^{\pm R}[f(\mathbf{x})] &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} f(\mathbf{x}) e^{\mp \mu \mathbf{u}^T \mathbf{x}} d\mathbf{x} = F^{\pm R}[\mathbf{u}] \Leftrightarrow \\ F^{\mp R}[F^{\pm R}[\mathbf{u}]] &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} F^{\pm R}[\mathbf{u}] e^{\pm \mu \mathbf{u}^T \mathbf{x}} d\mathbf{u} = f(\mathbf{x}) \end{aligned} \quad (13)$$

where $\mathbf{x} = (x, y)$, $\mathbf{u} = (u, v)$ (u, v are quaternion frequencies) and $\boldsymbol{\mu}$ represents a pure quaternion unit that defines the axis of the transformation. $\boldsymbol{\mu}$ is chosen to be $\boldsymbol{\mu} = (\mathbf{i} + \mathbf{j} + \mathbf{k})/\sqrt{3}$ in [9] for processing of natural RGB images as it is aligned with the grayline axis of the unit RGB colour cube.

4. Quaternion Potential Functions

4.1. Calculating Potential Functions using Quaternion Cross-Correlation

As already mentioned in Section 2.1, both the single node potential V_i and the pairwise potential $V_{ij}(l_i, l_j)$ are calculated by taking the SSD of the two image regions. Suppose that f and g are two image regions, the similarity between the shifted f and g within a region of interest in g is evaluated as [13]

$$SSD(\boldsymbol{\delta}) = \sum_{\mathbf{x}} (f(\boldsymbol{\chi} - \boldsymbol{\delta}) - g(\boldsymbol{\chi}))^2 w(\boldsymbol{\chi}) \quad (14)$$

where $\boldsymbol{\chi} = (x, y)$ indicates the coordinate of a pixel, $\boldsymbol{\delta}$ represents the shift and w is the binary mask of the same size as f and g that is only non-zero in the interested region (w is a matrix of all 1s when estimating the pairwise potential).

Equation (14) can be rewritten as

$$SSD(\boldsymbol{\delta}) = \sum_{\mathbf{x}} (f^2(\boldsymbol{\chi} - \boldsymbol{\delta})w(\boldsymbol{\chi})) - 2 \sum_{\mathbf{x}} (f(\boldsymbol{\chi} - \boldsymbol{\delta})g(\boldsymbol{\chi})w(\boldsymbol{\chi})) + \sum_{\mathbf{x}} (g^2(\boldsymbol{\chi})w(\boldsymbol{\chi})) \quad (15)$$

The last term of equation (15) is independent to $\boldsymbol{\delta}$, so we only need to calculate it once by summing the squares of all the values in g in the interested region (i.e. where $w(\boldsymbol{\chi}) = 1$). The first two terms can be evaluated by calculating the cross-correlation of f and g

$$\sum_{\mathbf{x}} (f^2(\boldsymbol{\chi} - \boldsymbol{\delta})w(\boldsymbol{\chi})) = [\overline{f^2}(\boldsymbol{\chi}) \star w(\boldsymbol{\chi})]_{\boldsymbol{\delta}} \quad (16)$$

$$\sum_{\mathbf{x}} (f(\boldsymbol{\chi} - \boldsymbol{\delta})g(\boldsymbol{\chi})w(\boldsymbol{\chi})) = [\overline{f}(\boldsymbol{\chi}) \star (g(\boldsymbol{\chi})w(\boldsymbol{\chi}))]_{\boldsymbol{\delta}} \quad (17)$$

where \star is the correlation operator and $\overline{f}(\boldsymbol{\chi})$ is the quaternion conjugate of f . Note that all the computations above are performed in the quaternion domain. Thus, we need to convert the input colour images into quaternion matrices using equation (12) and then calculate their cross-correlation.

The cross-correlation of two quaternion images f and g was originally extended from standard complex correlation using basic quaternion arithmetic [9, 11]

$$C(f, g) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} f(m, n) \overline{g(m-k, n-l)} \Rightarrow cr(k, l) . \quad (18)$$

$cr(k, l)$ is the resulting correlation surface as an $M \times N$ quaternion image. The SSD is then calculated by taking the modulus of the values in the correlation surface $cr(k, l)$.

4.2. QFT for Calculating Quaternion Cross-Correlation

It is impractical to evaluate the cross-correlation function in equation (18) directly due to the high computational cost ($O(N^4)$ for a $N \times N$ image [9]). Thus, a method of calculating the quaternion cross-correlation based on fast QFT was developed in [9].

$$\begin{aligned} C(f, g) &= F^{-R} \{ F^{+R} [\mathbf{u}] \overline{G_{\parallel}^{+R}[\mathbf{u}]} \} + F^{+R} \{ F^{+R} [\mathbf{u}] \overline{G_{\perp}^{+R}[\mathbf{u}]} \} \\ &= F^{\mp R} \{ F^{+R} [\mathbf{u}] \overline{G_{\parallel}^{\pm L}[\mathbf{u}]} + F^{-R} [\mathbf{u}] \overline{G_{\perp}^{\pm L}[\mathbf{u}]} \} \end{aligned} \quad (19)$$

where $\mathbf{u} = (u, v)$, $G(\mathbf{u}) = QFT\{g(m, n)\}$, $G_{\parallel}[\mathbf{u}] \parallel \mathbf{u}$ and $G_{\perp}[\mathbf{u}] \perp \mathbf{u}$

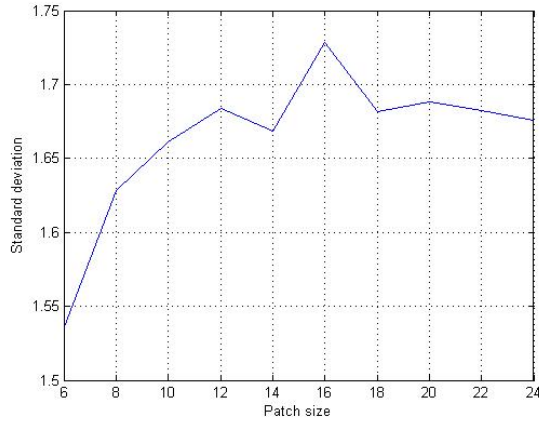
Using equation (19), we can rewrite equations (16) and (17) as

$$\begin{aligned} \overline{f^2}(\boldsymbol{\chi}) \star w(\boldsymbol{\chi}) &= F^{\mp R} \{ F^{+R} [\overline{f^2}(\boldsymbol{\chi})] \overline{G_{\parallel}^{\pm L}[w(\boldsymbol{\chi})]} \\ &+ F^{-R} [\overline{f^2}(\boldsymbol{\chi})] \overline{G_{\perp}^{\pm L}[w(\boldsymbol{\chi})]} \} \end{aligned} \quad (20)$$

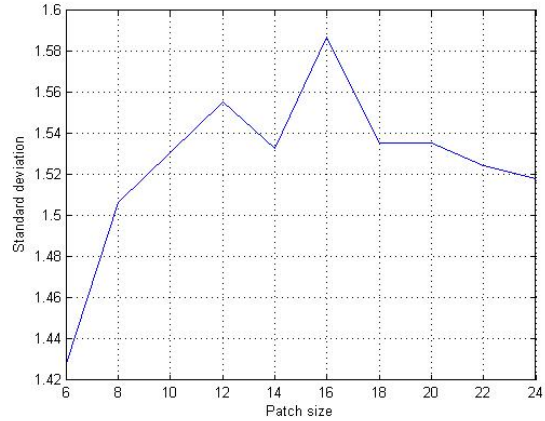
$$\begin{aligned} \overline{f}(\boldsymbol{\chi}) \star (g(\boldsymbol{\chi})w(\boldsymbol{\chi})) &= F^{\mp R} \{ F^{+R} [\overline{f}(\boldsymbol{\chi})] \overline{G_{\parallel}^{\pm L}[g(\boldsymbol{\chi})w(\boldsymbol{\chi})]} \\ &+ F^{-R} [\overline{f}(\boldsymbol{\chi})] \overline{G_{\perp}^{\pm L}[g(\boldsymbol{\chi})w(\boldsymbol{\chi})]} \} \end{aligned} \quad (21)$$

5. Experiments and Discussion

The proposed image completion method using hyper-complex potential functions was tested on a broad range of colour images including infrared images and images taken in both indoor and outdoor conditions. We compared the outputs obtained by performing the completion method on gray-scale images using the monochromatic model and colour images using quaternion potential functions to demonstrate the effectiveness of our method. For all experiments, we chose $w = h$ and $gap_x = gap_y = \frac{1}{2}h$ to get a good overlap region between two labels. The optimum size of patches and label pruning parameters used in the tests for gray-scale images were selected based on the automatic parametrisation method proposed in [7]. For colour images, we could calculate the standard deviation of the patch entropies for each channel and then take the average of the results in order to determine the optimum size. Figure 3 shows the plots of the standard deviation of the patch entropies versus patch size for Figure 5a in both gray-scale and colour. In both plots, we can see that patch size 16



(a) Gray-scale image



(b) colour image

Figure 3. Standard deviation of the patch entropies versus patch size.

gives the largest standard deviation for the patch entropies. Thus, the optimum size for the completion process for the given input image is 16 for both colour and monochromatic models.

For the far-infrared image (Figure 4), the results are similarly good for both gray-scale and colour images. This is not surprising as this far-infrared image essentially displays temperature values and thus only has one channel. The colour representation is merely a different way of visualising the different temperature values in the scene, although we treated the ‘colour’ representation for the sake of the experiments here as if it was a true 3-channel colour image. As the colour distribution of the background is nearly uniform in far-infrared images, the difference between the colour and gray-scale output is not significant. Another reason is that there are no complex background structures under the interested region in this example, so it is simpler to complete the image even in the gray-scale domain.

For the outdoor scene, there is a considerable difference in the quality of the gray-scale and the colour outputs (Figure 5). It is clear that the outputs for colour images (Figures 5d, 5h) are perceptually better than the gray-scale outputs (Figures 5c, 5g). In Figure 5c, we can still see implausible blocks when filling the lake surface and the fountain which are not present in Figure 5d. Although the water surface is completed well in both cases, the background region behind the boat is filled in a more plausible way in Figure 5h than in Figure 5g. Naturally, what happens in this case is that by using hypercomplex numbers to model colour images, the correlations across the channels are captured which provides more information to complete those images.

A number of further examples in Figure 6 show the effectiveness of our approach in completing images in the presence of a variety of structures in the background. These

input images were taken indoors, containing complex structures like tables, chairs and other objects. In all cases, these images were convincingly reconstructed by applying the hypercomplex colour model (Figures 6d, 6h). These results outperformed the ones for gray scale images (Figures 6c, 6g).

6. Conclusions

In this paper, we have presented a method of using hypercomplex numbers, also known as quaternions, and their Fourier transforms to calculate the similarity of colour image patches and then estimate the potential functions of an image completion algorithm using MRFs and BP. This method provides a systematic, unified way of solving colour image completion instead of processing each channel separately. By using quaternions to model RGB pixels, we can improve the performance of the completion algorithm. The proposed approach can also be used in other problems that need to calculate the similarity of image regions including colour image alignment, object recognition in colour images, and super-resolution. Results comparing this approach to the completion method for monochromatic image show the effectiveness of our approach. In future work, the same quaternion idea can be applied to the automatic parametrisation process to achieve a better estimation of parameters for the image completion problem.

References

- [1] M. Bertalmio, G.Sapiro, V. Caselles, and C. Ballester. Image Inpainting. In *Proc. ACM SIGGRAPH 2000*, pages 417–424, 2000.



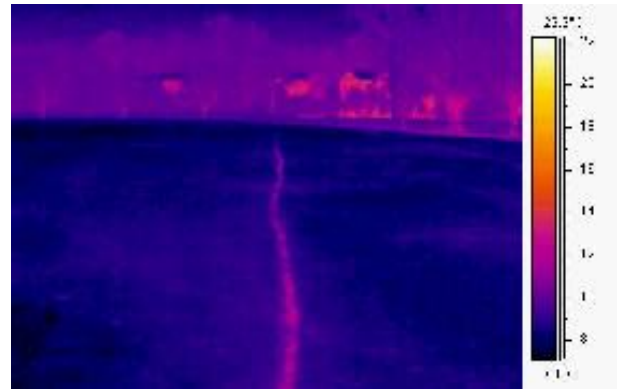
(a) Infrared image



(b) Mask



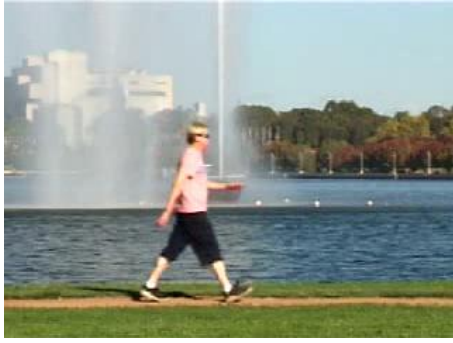
(c) Gray-scale output



(d) colour output

Figure 4. Image completion results for the far-infrared input image.

- [2] A. Efros and T. Leung. Texture synthesis by non-parametric sampling. *International Journal of Computer Vision*, pages 1033–1038, 1999.
- [3] T. Ell. *Hypercomplex spectral transforms*. PhD. dissertation, Univ. Minnesota, Minneapolis, 1992.
- [4] P. Felzenszwalb and D. Huttenlocher. Efficient Belief Propagation for Early Vision. *International Journal of Computer Vision*, 70(1):41–54, 2006.
- [5] W. Freeman, E. Pasztor, and O. Carmichael. Learning Low-level vision. *International Journal of Computer Vision*, 40(1):25–47, 2000.
- [6] W. Hamilton. *Elements of Quaternions*. London, U.K.: Longmans, Green, 1866.
- [7] H. Ho and R. Goecke. Automatic Parametrisation for an Image Completion Method Based on Markov Random Fields. In *2007 IEEE International Conference on Image Processing*, Sept. 2007. (accepted).
- [8] N. Komodakis and G. Tziritas. Image Completion Using Global Optimization. In *Proc. CVPR2006*, volume 1, pages 442–452, New York, USA, June 2006.
- [9] C. Moxey, S. Sangwine, and T. Ell. Hypercomplex correlation techniques for vector images. *IEEE Transactions on Signal Processing*, 51(7):1941–1953, July 2003.
- [10] J. Protilla and E. Simoncelli. A parametric texture model based on joint statistics of complex wavelet coefficients. *International Journal of Computer Vision*, 40(1):49–70, 2000.
- [11] S. Sangwine and T. Ell. Color image filters based on hypercomplex convolution. In *Proc. Inst. Elect. Eng. Vision, Image and Signal Processing*, volume 147(2), pages 89–93, 2000.
- [12] J. Sun, L. Yuan, J. Jia, and H.-Y. Shum. Image completion with structure propagation. In *Proceedings of the ACM SIGGRAPH*, volume 24, pages 861–868, 2005.
- [13] A. Wong, W. Bishop, and J. Orchard. Efficient Multi-Modal Least-Squares Alignment of Medical Images Using Quasi-Orientation Maps. In *Proc. IPCV 2006*, pages 66–73, 2006.
- [14] C. Xie and B. V. Kumar. Quaternion correlation filters for color face recognition. In *Proceedings of SPIE*, volume 5681, pages 486–494, 2005.



(a)



(b)



(c)



(d)



(e)



(f)



(g)



(h)

Figure 5. Image completion results for outdoor scene. (a), (e): Background images, (b), (f): Binary masks, (c), (g): Results for gray-scale images, (d), (h): Results for colour images.



(a)



(b)



(c)



(d)



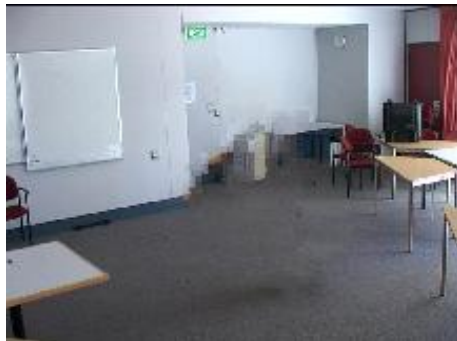
(e)



(f)



(g)



(h)

Figure 6. Image completion results for indoor scene. (a), (e): Background images, (b), (f): Binary masks, (c), (g): Results for gray-scale images, (d), (h): Results for colour images.