Image Renaissance Using Discrete Optimization and the α-Expansion Algorithm

Cédric Allène
Nikos Paragios

Research Report 05-11
June 2005

CERTIS, ENPC,
77455 Marne la Vallee, France,

Image Renaissance Using Discrete Optimization and the α-Expansion Algorithm

Renaissance d'Image utilisant l'Optimisation Discrète et l'Algorithme d'α-Expansion

Cédric Allène[1,2]
Nikos Paragios[2]

[1]A^{2}SI, ESIEE, 93162 Noisy-le-Grand, France, http://www.esiee.fr/~info/a2si/
[2]CERTIS, ENPC, 77455 Marne la Vallee, France, http://www.enpc.fr/certis/

Abstract

In this paper we propose a novel technique that addresses image renaissance through a ”multi-level” graph-based matching process. To this end, numerous image patches that do present similarities with the local content around the missing part are considered. The selection of these patches is done through a particle filter method to address the task of hypotheses evaluation. These patches are positioned on top of missing segment, ordered depending on their similarity weight, and form in some fashion a multi-layered graph over time. Markov Random Fields are used to formalize inpainting as a labelling estimation problem while a combinatorial approach is used to recover the optimal combination of patches to complete the missing structure. The min-cut max-flow algorithm within the α-expansion process is used to determine the optimal cut that, in an implicit fashion, completes the missing image structure. Promising results in image and texture completion demonstrate the potentials of the proposed method.

Résumé

Ce document propose une nouvelle technique effectuant la renaissance d'image grâce à un procédé de comparaison par graphes successifs. Dans ce but, nous considérons des régions de l'image qui ont des ressemblances avec les données avoisinant les zones manquantes. La sélection de ces régions se fait grâce à la méthode du filtre à particules pour calculer la mesure de ressemblance. Ces régions sont positionnées par-dessus les parties manquantes, ordonnés selon le poids de leur similarité, formant une sorte de graphe à plusieurs niveaux dans le temps. Les "Markov Random Fields" sont utilisés pour représenter l'inpainting comme un problème d'estimation d'étiquetage tandis qu'une approche combinatoire est utilisée pour retrouver la combinaison optimale de régions complétant la partie manquante. L'algorithme de min-cut/max-flow dans le procédé d'α-expansion sert à déterminer la coupe optimale qui, de façon implicite, complète la structure manquante de l'image. Des résultats prometteurs dans la complétion d'image et de tectures démontrent le potentiel de la méthode proposée.

Contents

1 Introduction

Image inpainting [3], often called retouching, consists of completing missing or damaged parts of an image. Such a demand was first addressed in painting restoration and has spread in other domains of computational vision like photos [25], films [23] and recently 3D volumes that consist of appearance and structure [2]. Prior art to image completion consists of variational and statistical methods.

Variational methods are mostly inspired by the Euler elastica model [27], a mathematical formulation for recovering missing geometric planes according to the principle of good continuation [22] given a start, an end point and their normals. In [26] image inpainting was also called image disocclusion. Their approach aim to decompose the image in isophotes of constant appearance and then to complete these isophotes within the "occluded" part according to the elastica model [27]. Such a concept was explored in its geometric form in [3] according to a PDE that was propagating information along the image isophotes. Similar concept was considered in [9] where a curvature-driven diffusion approach was used to complete missing image structure. Joint interpolation of image and vector fields within a variational formulation was a further attempt to explore the elastica formulation while relaxing the boundary conditions [1] while several variants of the total variation minimization algorithm were considered to address inpainting [8]. Last, but not least, a variation of the Mumford-Shah framework was proposed in [14] to complete the missing structure. While these methods are quite popular and have an outstanding performance on filling micro-gaps, one can claim that dealing with texture[3] as well as with important missing image regions are their most notable limitations.

Statistical methods were made popular in texture synthesis, a problem that is somehow related with the inpainting one. The central idea is to reproduce a pattern at a local scale (missing gaps) that is stationary [28]. This concept was formulated in a mathematical fashion in [13] within Markov Random Fields [16] and spread across the vision and graphics domain. One can refer to methods that aim to reconstruct textures from sample images through the selection of an appropriate set of filter operators [29], or techniques with aim to replace the missing part with an "artificial" texture like the ones introduced in [21]. Last, but not least one can refer to recent approaches where texture synthesis consists of stitching together blocks of existing sample texture a method that was explored in [12, 24] within graph-based combinatorial optimization. Opposite to variational methods one can claim that statisti-

[3]That problem was addressed in some partial fashion in [4].

cal methods can do a much better job when dealing with texture and have a better performance when dealing with macro-gaps under the assumption that the correct scale was considered.

(a) (b) (c)

Figure 1: *Reconstruction of muppet PPD: (a) input image, (b) optimal partition, (b) reconstructed image.*

In this paper, we address the problem of image renaissance through a graph-based approach, that aims to combine variational and statistical methods that is based on the concept of progressive stitching within a temporal process as shown in [Fig. (1)].

To this end, we assume that one can find in the image portions of the missing information. Once candidates have been determined thanks to a selection process of seeds indicating the origin of an image area similar to data at the bounadry of the missing part, we reformulate image renaissance as a problem of min-cut within the graph. To this end, a MRF-based cost function is presented that accounts for the similarity of the existing segments and the ones to be added that is evolving upon the completion of the missing structure. Furthermore, within a temporal process one can introduce metrics that do account for the distance from the boundaries of the inpainted region and in a non-explicit fashion we do propagate the information along the isophotes. Towards addressing various scales and orientations, a multi-source approach is considered where a substantial number of candidate image segments are super-imposed to the one to be inpainted. The min-cut is recovered through max flow problem principle according to the α-expansion algorithm.

The reminder of this paper is organized according to the following fashion. In section 2 we introduce the theoretical concept of our approach. The selection process for the candidate seeds is presented in section 3 while in section 4 we present the considered graph-based optimization principles. while discussion and experimental results are part of section 5.

2 Image Completion

Let us consider an image $\mathcal{I}' : \Omega \to \mathcal{R}$, that is a replica of the original image $\mathcal{I} : \Omega \to \mathcal{R}$ with the exception of a mask $\mathcal{S}$ that refers to the image segments that are not present. The task of inpainting consists of creating a new image $\mathcal{F}$ such that

$$\mathcal{F}(\mathbf{x}) = \begin{cases} \mathcal{I}'(\mathbf{x}), \mathbf{x} \in \Omega - \mathcal{S} \\ \mathcal{I}'(\mathbf{x}), \mathbf{x} \in \mathcal{S} \end{cases}$$

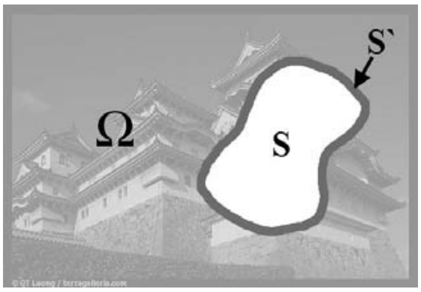

Figure 2: Concept of Image Completion through successive patch matching.

Without loss of generality - since the image $\mathcal{I}$ is not available - one can assume that the missing part $\mathcal{F}(\mathcal{S})$ can be reconstructed through other present image segments;

$$\mathcal{F}(\mathbf{x}) = \begin{cases} \mathcal{I}'(\mathbf{x}), & \mathbf{x} \in \Omega - \mathcal{S} \\ \mathcal{I}'(\mathbf{x}'), & \mathbf{x} \in \mathcal{S}, \mathbf{x}' \in \Omega - \mathcal{S} : \mathcal{I}'(\mathbf{x}') = \mathcal{I}(\mathbf{x}) \end{cases}$$

that is the core assumption of our method. Let us consider n image segments

$$\{\mathcal{L}_1, \mathcal{L}_2, ..., \mathcal{L}_n\} : \mathcal{L}_i \in \Omega - \mathcal{S}$$

randomly positioned over $\mathcal{S}$. Then the problem of inpainting consists of selecting for every pixel of $\mathcal{S}$, the value among these n possible ones that best approximates the original data. One can see such a task in the form of a labelling problem, where to the pixel $\mathbf{x}$ the label $\omega(\mathbf{x}) \in [1, n]$ is attributed reflecting to: $\mathcal{F}(\mathbf{x}) = \mathcal{L}_{\omega(\mathbf{x})}(\mathbf{x})$.

However, in the absence of constraints within the inpainted region $\mathcal{S}$ one can relax the constraint to account for the image properties on the borders of $\mathcal{S}$ leading to a new inpainted region $\mathcal{S}'$ as shown in [Fig. (2)] where constraints are present. Then, towards solving the labelling problem the following objective function can be considered:

$$E(\omega) = \sum_{\mathbf{x} \in \mathcal{S}' - \mathcal{S}} \left(\mathcal{L}_{\omega(\mathbf{x})}(\mathbf{x}) - \mathcal{I}'(\mathbf{x})\right)^2$$

where the distance between the new inpainted image and the existing observations is minimal. One can consider more advanced error metrics that perform an evaluation at a local region/neighbourhood [$\mathcal{N}(\mathbf{x})$] level rather than at the pixel level, or

$$E(\omega) = \sum_{\mathbf{x} \in \mathcal{S}' - \mathcal{S}} \rho\left(\mathcal{L}_{\omega(\mathbf{x})}(\mathcal{N}(\mathbf{x})), \mathcal{I}'(\mathcal{N}(\mathbf{x}))\right)$$

where $\rho()$, often called the local potential function, can be for example the correlation of intensities between the two image patches. Within our approach we assume a local normalized SSD score in the vertical and the horizontal direction pointed by a Gaussian distribution:

$$\rho\left(\mathcal{L}_{\omega(\mathbf{x})}(\mathbf{x}), \mathcal{I}'(\mathbf{x})\right) = \frac{1}{\mathrm{Z}} \sum_{\mathbf{m}=-\mathcal{W}}^{\mathcal{W}} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{|\mathbf{m}|^2}{2\sigma^2}\right) \left|\mathcal{L}_{\omega(\mathbf{x})}(\mathbf{x}+\mathbf{m}) - \mathcal{I}'(\mathbf{x}+\mathbf{m})\right|$$

where Z is a normalization factor. However, visual discontinuities are quite noticeable from biological vision systems while important sporadic changes of small magnitude within uniform regions often do not attract the attention of the humans.

The use of the image derivatives as well as the ones of the patch under consideration could be a considered as support layers on this constraint. To this end, in [24] a term that is inversely proportional to the norm of the gradient was considered within the context of texture synthesis,

$$E(\omega) = \sum_{\mathbf{x} \in \mathcal{S}' - \mathcal{S}} \exp\left\{\frac{\rho\left(\mathcal{L}_{\omega(\mathbf{x})}(\mathbf{x}), \mathcal{I}'(\mathbf{x})\right)}{|\nabla \mathcal{I}'(\mathbf{x})|^2 + |\nabla \mathcal{L}_{\omega(\mathbf{x})}(\mathbf{x})|^2}\right\}$$

that will make pixels with substantial derivatives more significant in the reconstruction process.

Such a method will be able to reconstruct the image at the pixel level while preserving discontinuities through an independent decision process according to the similarity between the observed image and the candidate patches. Such an independent process will form several discontinuities that will be quite disrupting to the human eye and will violate the condition that images are assumed to be consistent at a local scale.

Such a limitation is often addressed using local smoothness constraints on the label domain, that consists of saying that neighbourhood pixels should have about the same label;

$$E(\omega) = \sum_{\mathbf{x} \in \mathcal{S}' - \mathcal{S}} \left[\sum_{\mathbf{x} \in \mathcal{N}(\mathbf{x})} \mathcal{V}(\omega(\mathbf{x}), \omega(\mathbf{y}))\, d\mathbf{y}\right]$$

where the function $\mathcal{V}$ in the most general case has the following form

$$\mathcal{V}(\omega(\mathbf{x}), \omega(\mathbf{y})) = \begin{cases} +\alpha_{diff}, & \omega(\mathbf{x}) \neq \omega(\mathbf{y}) \\ 0, & \omega(\mathbf{x}) = \omega(\mathbf{y}) \end{cases}$$

with $\alpha_{diff} > 0$. It is important to note that such a term imposes smoothness on the label space of the reconstructed image that is not equivalent with smoothness in the image itself.

One can now consider the image potential and the smoothness term in their discrete form - the most common MRF-based formulation - and use them to address inpainting once the smoothness constraint has been relaxed within the $\mathcal{S}'$ according to:

$$E(\omega) = \sum_{\mathbf{x} \in \mathcal{S}' - \mathcal{S}} \exp \left\{ \frac{\rho\left(\mathcal{L}_{\omega(\mathbf{x})}(\mathbf{x}), \mathcal{I}'(\mathbf{x})\right)}{|\nabla \mathcal{I}'(\mathbf{x})|^2 + |\nabla \mathcal{L}_{\omega(\mathbf{x})}(\mathbf{x})|^2} \right\} + \beta \sum_{\mathbf{x} \in \mathcal{S}'} \left[\sum_{\mathbf{x} \in \mathcal{N}(\mathbf{x})} \mathcal{V}(\omega\left(\mathbf{x}\right), \omega(\mathbf{y}))\, d\mathbf{y} \right]$$

where β is a constant, balancing the contribution of these two terms. Such an objective function aims to create a partition on the labelling space such that the existing part of the image remains mostly the same while for the region to be inpainted is a smooth continuation on the labelling space of the one of the image (that is conceptually different from the essence of variational methods). One can seek the lowest -sub-optimal- potential of the discrete form of function using several techniques of various complexity like the iterated conditional modes [5], the highest confidence first [10], the mean-field and simulated annealing [16] and the min-cut max flow approach [7]. The most important limitation of this approach is that inpainting is done through a stitching process on the boundaries of the inpainted region while poorly smoothness constraints are used to fill in the missing information.

In order to overcome this limitation let us consider a temporal process where ω is a function of time $[\omega^0, ..., \omega^t]$ where the information is propagated in a progressive fashion $[\mathcal{I}' = \mathcal{F}^0, ..., \mathcal{F}^t]$ and as $t \to \infty$ the region to be inpainted vanishes $[\mathcal{S} = \mathcal{S}^0 \supset ... \supset \mathcal{S}^t = \emptyset]$. Then one can consider a procedure that optimizes the same cost function at various time instants that upon converge performs inpainting in an optimal fashion:

$$\begin{aligned}
&\lim_{t\to\infty} E(\omega^t) = \\
&\alpha \sum_{\mathbf{x}\in\mathcal{S}'-\mathcal{S}^0} \exp\left\{\frac{\rho\left(\mathcal{L}_{\omega(\mathbf{x})}(\mathbf{x}), \mathcal{I}'(\mathbf{x})\right)}{|\nabla\mathcal{I}'(\mathbf{x})|^2 + |\nabla\mathcal{L}_{\omega(\mathbf{x})}(\mathbf{x})|^2}\right\} \\
&+\beta \sum_{\mathbf{x}\in\mathcal{S}'-\mathcal{S}^{t-1}} \exp\left\{\frac{\rho\left(\mathcal{L}_{\omega(\mathbf{x})}(\mathbf{x}), \mathcal{F}^{t-1}(\mathbf{x})\right)}{|\nabla\mathcal{I}'(\mathbf{x})|^2 + |\nabla\mathcal{L}_{\omega(\mathbf{x})}(\mathbf{x})|^2}\right\} \\
&+\gamma \sum_{\mathbf{x}\in\mathcal{S}'} \left[\sum_{\mathbf{x}\in\mathcal{N}(\mathbf{x})} \mathcal{V}(\omega^t(\mathbf{x}), \omega^t(\mathbf{y}))\, d\mathbf{y}\right]
\end{aligned}$$

with $\alpha \gg \beta > \gamma$. One can interpret these terms in the following fashion. The first term constraints the superimposed segments to match with the image content on the borders of the region to be inpainted, a rather hard constraint that is reflected on the selection of the α coefficient. The second term, updates in a progressive fashion the content within the inpainted region while the third term imposes smoothness on the label space of the reconstructing image that is not equivalent with smoothness in the image itself.

Despite the theoretical advantages of such an approach, in practice it cannot be considered. In order to complete the missing structure several patches are to be considered in a random fashion. Consequently, the number of labels is substantial > 1000. Therefore even the optimization of such a cost function at a single scale is quite expensive unless efficient sub-optimal techniques are considered.

One can relax the constraint of time and assume that the series

$$[\mathcal{S} = \mathcal{S}^0 \supset ... \supset \mathcal{S}^t = \emptyset]$$

can be constructed according to a linear function of the Euclidean distance from borders of the inpainted region. In other words one can consider isophotes in the Euclidean distance space and not in the image itself that can be pre-computed in advance and simulate time according to this distance;

$$\begin{aligned}
E(\omega) = \alpha &\sum_{\mathbf{x}\in\mathcal{S}'-\mathcal{S}} \exp\left\{\frac{\rho\left(\mathcal{L}_{\omega(\mathbf{x})}(\mathbf{x}), \mathcal{I}'(\mathbf{x})\right)}{|\nabla\mathcal{I}'(\mathbf{x})|^2 + |\nabla\mathcal{L}_{\omega(\mathbf{x})}(\mathbf{x})|^2}\right\} \\
+\beta &\sum_{\mathbf{x}\in\mathcal{S}} \exp\left\{\frac{\sigma^2}{\mathcal{D}^2(\mathbf{x}, \partial\mathcal{S})} \frac{\rho\left(\mathcal{L}_{\omega(\mathbf{x})}(\mathbf{x}), \mathcal{F}(\mathbf{x})\right)}{|\nabla\mathcal{I}'(\mathbf{x})|^2 + |\nabla\mathcal{L}_{\omega(\mathbf{x})}(\mathbf{x})|^2}\right\} \\
&+\gamma \sum_{\mathbf{x}\in\mathcal{S}'} \left[\sum_{\mathbf{x}\in\mathcal{N}(\mathbf{x})} \mathcal{V}(\omega(\mathbf{x}), \omega(\mathbf{y}))\, d\mathbf{y}\right]
\end{aligned}$$

where $\mathcal{D}(\mathbf{x}, \partial\mathcal{S})$ is the minimum Euclidean distance between the pixel $\mathbf{x}$ and the interface that delineates the boundaries of the region to be inpainted, σ is a constant parameter sampling the importance of progressive reconstruction and as earlier explained $\alpha \gg \beta > \gamma$.

The next issue to be addressed now is the selection and the position of candidate seeds that has to be as small as possible, resulting to plausible solution on the labelling process. Random selection of image patches is a straightforward solution to the problem that will though make the approach quite inefficient.

3 On the Selection of Candidate Seeds

We call seeds the positions in the image whom neighbourhood has good similarities with the neighbourhood of a point in the boundary of the part to be inpainted. One can consider the problem of seed extraction as a tracking problem in the image where given a starting position, the objective is to recover an image region that can be used to replace the missing segment. The statistical interpretation of such an objective refers to the introduction of a probability density function (pdf) that uses previous states to predict possible new positions for the added seeds, while image features are used to evaluate the quality of these predictions.

Let us consider a state vector $\omega = (\mathbf{x}, \mathbf{l}_x, \mathbf{l}_y)$ that describes a rectangle $(\mathbf{l}_x, \mathbf{l}_y)$ that is centred at $\mathbf{x}$. Particle filters [20] are sequential Monte-Carlo techniques that can be used to estimate the Bayesian posterior probability density function with a set of samples [18] ω_t, conditional to observations from time 1 to time t $\mathbf{z}_{1:t}$: $p(\omega_t|\mathbf{z}_{1:t})$.

In terms of a mathematical formulation, such a method approximates the posterior pdf by M random measures $\{\mathbf{z}_t^m, m = 1, M\}$ associated to M weights $\{w_t^m, m = 1, M\}$, such that

$$p(\omega_t|\mathbf{z}_{1:t}) \approx \sum_{m=1}^{M} w_t^m \delta(\omega_t - \omega_t^m).$$

where each weight w_t^m reflects the importance of the sample ω_t^m in the pdf.

The samples ω_t^m are drawn using the principle of *Importance Density* [19], of pdf $q(\omega_t|\omega_{1:t}^m, \mathbf{z}_t)$, and it is shown that their weights w_t^m are updated according to

$$w_t^m \propto w_{t-1}^m \frac{p(\mathbf{z}_t|\omega_t^m)p(\omega_t^m|\omega_{t-1}^m)}{q(\omega_t^m|\omega_{t-1}^m, \mathbf{z}_t)}.$$

Figure 3: *Several steps of particule filter seeds selection (blue rectangles are patches tested to overlap the green rectangles area around the boundary of the missing part.*

Once a set of samples has been drawn, $p(\omega_t^m|\omega_{t-1}^m, \mathbf{z}_t)$ can be computed out of the observation $\mathbf{z}_t$ for each sample, and the estimation of the posteriori pdf can be sequentially updated.

The *Importance Density* $q(x_t|x_{1:t}^m, z_t)$ is chosen to be equal to $p(\omega_t^m|\omega_{t-1}^m)$, which reduces the previous equation to

$$w_t^m \propto w_{t-1}^m p(\mathbf{z}_t|\omega_t^m).$$

This choice is not only the simplest to implement, but it has also been proved [11] it minimizes the variance of true weights $\{w^m\}$ conditional to $\left\{\omega_{t-1}^m\right\}$ $\{\mathbf{z}_t\}$. After few time steps, all weights but few become nearly null. Therefore, it is necessary to perform a re-sampling. We chose to use the *Systematic Importance Re-sampling* algorithm [18], which draws new samples from the posterior pdf $p(\omega_t|\mathbf{z}_{1:t})$.

The last issue to be dealt with is the definition of a measure between a prediction and the actual observation, used to compute the conditional probability $p(\mathbf{z}_t|\omega_t^m)$. Evaluation of hypotheses is a rather critical condition in particle filters. To this end, given the origin of the filter (distribution on the borders of the inpainted region), one would like to evaluate the fit of the particle under consideration. Such a task is equivalent of measuring dissimilarities between image patches, or distributions.

In practice, given the region to be inpainted and a uniform sampling rule along its borders a number of regions are considered centred at the boundary points. One then applies a random perturbation on the position of the centre of the rectangles. Given a new position and the new characteristics the quadratic distance between the new rectangle and the one of the origin is computed and used to update the weight of the particle. These weights guide the re-sampling process (more samples for particles with important weights) as well as the random perturbations (inversely proportional to these weights). Upon certain number of iterations, a fraction of the particles having the best weights is retained and superimposed to the region to be inpainted.

One now can consider the optimization of such a cost function to address inpainting. Now that candidate seeds have been determined, we have to find which pixel for each corresponding seed position will finally be pasted in the output reconstructed area. The objective function will be minimized through a combinatorial approach based on the graph-cut framework [7].

4 α-Expansion Combinatorial Optimization

Let us first introduce some basic graph terminology to facilitate the introduction of the method.

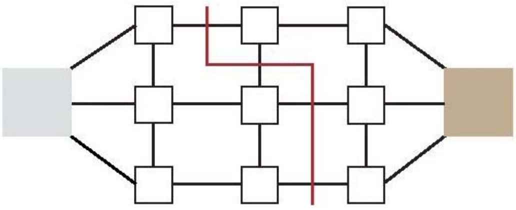

Figure 4: Example of graph with a cut.

$\mathcal{G}$=(ν, ε) is considered to be a graph that consists of a set of nodes ν and a set of directed edges ε which connect the nodes to each other. Furthermore, one can assume by construction that such a graph contains two special *terminal* nodes, the *source*, s, and the *sink*, t. All edges in the graph are assigned some non-negative weight or cost $w(p,q)$. The cost of the directed edge (p,q) may differ from its reverse edge (q,p). There are two types of edges: t-links, which are edges connecting a non-terminal node with a terminal node, and n-links, which are edges connecting two non-terminal nodes.

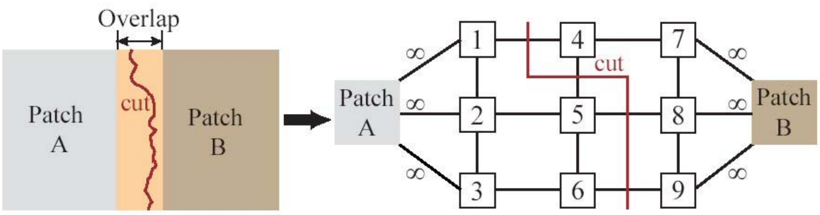

Figure 5: Graph construction and cut interpretation for a 3*3 pixels overlapping area (figure from [24]).

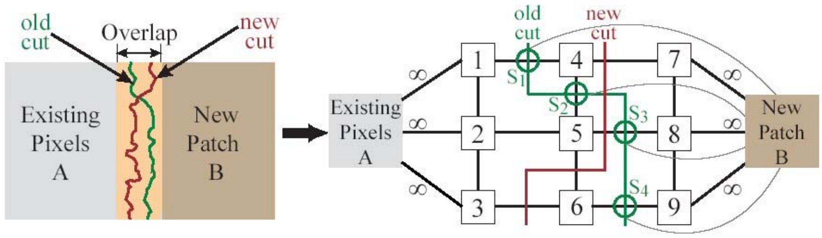

Figure 6: Graph construction and cut interpretation over an old seam for a 3*3 pixels overlaping area (figure from [24]).

A s/t cut C, often just called cut, on a graph $\mathcal{G}$ is a partitioning of the nodes in the graph into two subsets $\mathcal{S}$ and $\mathcal{T}$ with s in $\mathcal{S}$ and t in $\mathcal{T}$ as shown in [Fig. (4)]. The cost of a cut C=$\{\mathcal{S}$,$\mathcal{T}\}$ is the sum of the costs of the "boundary" edges (p,q) such that $p \in \mathcal{S}$ and $q \in \mathcal{T}$. The *minimum cut* problem on a graph consists in finding the cut that has the minimum cost among all possible cuts. It corresponds in finding the *maximum flow* going from the source s to the sink t [15]. The min-cut/max flow principle states that the maximum flow from s to t saturates a set of edges in the graph creating a partition of the nodes $\{\mathcal{S}$,$\mathcal{T}\}$ that corresponds to a minimum cut. One can find in the literature [6] numerous polynomial time algorithms exist solving the min-cut/max-flow problem. "Augmenting path" methods like the one considered in this paper [7] and "push-relabel" methods [17] is the most prominent categorization of these algorithms.

The α-expansion algorithm [7] consists of an iterative process that often converges to a local minimum through successive bin cuts. Within each step, among the n possible labels (which represent in fact copies oh the input image with an offset corresponding to the one given in the seed selection), the one corresponding to the seed having the best weight is selected and a cut that is optimal between this label and the actual output is computed. To this end, a graph is constructed in the following fashion (as shown in [Fig. 5]):
- one node is created for each pixel position in the overlaping area between the shifted image and the actual output,
- two extra nodes, the terminals, representing respectively the shifted image and the actual output
- for each couple of nodes which are neighbours in 4-connexity, an undirected n-link (which is equivalent to two directed edges in opposite directions) is added between the two nodes and weighted with the cost function $\rho()$ introduced in section 2 (where $\mathcal{N}(\mathbf{x})$ is a few pixels on the same line than the couple considered)

- undirected t-links are added for each pixel on the border of the overlapping area, connecting the corresponding node to the terminal representing the type of data next to the area
- extra nodes are inserted, with two n-links and one t-link at each "old boundary" edge of the partition, instead of the classical edge explained before, as shown in [Fig. (6,7)]

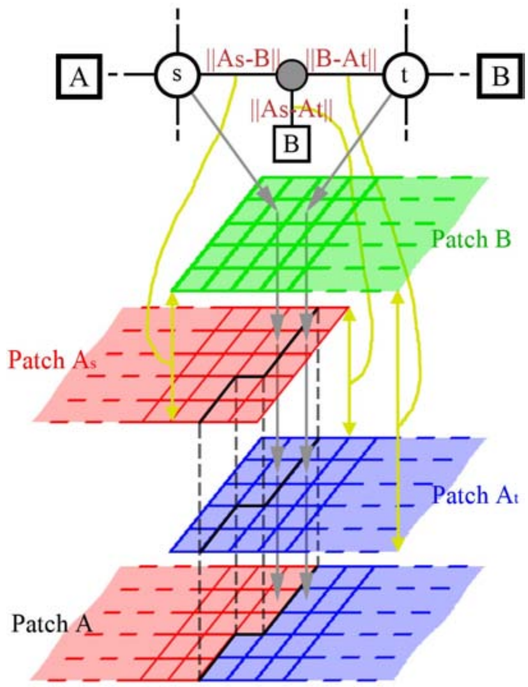

Figure 7: Graph construction and cut interpretation over an already computed image A (figure from [24]).

Once the graph is constructed, the graphcut algorithm is applied. The cut give us the boundary which offers the best transition between the two patches. So the result gives for each pixel if we have to copy the pixel corresponding to the label we are trying to paste or if we keep the actual one. If the pixel from the zone to be inpainted is still empty, the one from the label is directly copied. Then, the output image is updated between successive cuts. Such a sequential application of the method, is shown in [Fig. (8,9)].

Figure 8: *Reconstruction of (1) Monty, (2) Charete, (3) Japanese animation image; (a) input image, (b) Optimal partition, (b) reconstructed image (the images are courtecy [3]).*

5 Discussion

In this paper we have introduced a graph-based combinatorial approach to image renaissance. Such a method is based on the concept of stitching, where multiple candidates patches are superimposed to the missing segment forming a multi-source labelling problem solvable by graphs. Distances from the borders of the inpainted region are considered to introduce the notion of time where image completion is done in a progressive fashion. Segments of the inpainted region to which image information is present in their local neighbourhood are reconstructed first, moving gradually to areas where the inference problem is completely ill-posed. Very promising experimental results demonstrate the potentials of our approach as shown in [Fig. (8,9)].

Computational complexity is the main limitation of the proposed framework. The cost is polynomial to the number of candidate seeds and therefore particular attention is paid on the selection of seeds. Extended neighbourhood systems can also be considered to further impose continuity on the reconstruction process while the use of Graphic Processing Units to accelerate the process is also a promising direction.

(1)

(2)

(a) (b) (c)

Figure 9: *Reconstruction of (1) photo, (2) Micro; (a) input image, (b) Optimal partition, (b) reconstructed image (the images are courtecy [3]).*

References

[1] C. Ballester, M. Bertalmio, V. Caselles, G. Sapiro, and J. Verdera. Filling-in by Joint Interpolation of Vector Fields and Grey Levels. *IEEE Transactions on Image Processing*, 10:1200–1211, 2001.

[2] C. Ballester, V. Caselles, and J. Verdera. Disocclusion by Joint Interpolation of Vector Fields and Gray Levels. *Multiscale Modeling and Simulation*, 2:80–123, 2003.

[3] M. Bertalmio, G. Sapiro, L.-T. Cheng, and S. Osher. Image Inpainting. In *ACM SIGGRAPH*, pages 417–424, 2000.

[4] M. Bertalmio, L. Vese, G. Sapiro, and S. Osher. Simultaneous structure and texture image inpainting. *IEEE-TIP*, 12(8):882–889, 2003.

[5] J. Besag. On the statistical analysis of dirty images. *Journal of Royal Statistics Society*, 48:259–302, 1986.

[6] Y. Boykov and V. Kolmogorov. An Experimental Comparison of Min-Cut/Max-Flow Algorithms for Energy Minimization in Vision. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26:1124–1137, 2004.

[7] Y. Boykov, O. Veksler, and R. Zabih. Fast Approximate Energy Minimization via Graph Cuts. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23:1222–1239, 2001.

[8] T. Chan and J. Shen. Mathematical Models of Local Non-texture Inpaintings. *SIAM Journal on Applied Mathematics*, 62:1019–1043, 2001.

[9] T. Chan and J. Shen. Non-texture Inpaintings by Curvature-driven Diffusions (CDD). *Journal of Visual Communication and Image Representation*, 12(4):436–449, 2001.

[10] P. Chou and C. Brown. The theory and practice of bayesian image labeling. *International Journal of Computer Vision*, 4:185–210, 1990.

[11] A. Doucet, J. de Freitas, and N. Gordon. *Sequential Monte Carlo Methods in Practice*. Springer-Verlag, New York, 2001.

[12] A. Efros and W. Freeman. Image Quilting for Texture Synthesis and Transfer. In *Proc. SIGGRAPH, ACM Press*, pages 341–346, 2001.

[13] A. Efros and T. Leung. Texture synthesis by non-parametric sampling. In *IEEE International Conference on Computer Vision*, pages 1033–1038, 1999.

[14] S. Esedoglu and J. Shen. Digital Inpainting Based on the Mumford-Shah-Euler Image Model. *European J. Appl. Math.*, 13:353–370, 2002.

[15] L. Ford and D. Fulkerson. *Flows in Networds*. Princeton University Press, 1962.

[16] S. Geman and D. Geman. Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6:721–741, 1984.

[17] A. Goldberg and R. Tarjan. A New Approach to the Maximum Flow Problem. *Journal of the Association for Computing Machinery*, 35:921–940, 1988.

[18] N. Gordon. Novel Approach to Nonlinear/Non-Gaussian Bayesian State Estimation. *IEE Proceedings*, 140:107–113, 1993.

[19] N. Gordon. On Sequential Monte Carlo Sampling Methods for Bayesian Filtering. *Statistics and Computing*, 10:197–208, 2000.

[20] N. Gordon. A Tutorial on Particle Filters for On-line Non-linear/Non-Gaussian Bayesian Tracking. *IEEE Transactions on Signal Processing*, 50:174–188, 2002.

[21] D. Heeger and J. Bergen. Pyramid-Based Texture Analysis/Synthesis. In *Proceedings of ACM SIGGRAPH 95, ACM Press*, pages 229–238, 1995.

[22] G. Kanizsa. *Gramática de la Visión*. Paidos, 1986.

[23] A. Kokaram. On missing data treatment for degraded video and film archives: a survey and a new bayesian approach. *IEEE Transactions on Image Processing*, 13:397–415, 2004.

[24] V. Kwatra, A. Schodl, I. Essa, G. Turk, and a. Bobick. Graphcut Textures: Image and Video Synthesis Using Graph Cuts. In *Proceedings of SIGGRAPH*, 2003.

[25] S. Masnou. *Filtrage et Desocclusion d'Images par Méthodes d'Ensembles de Niveau*. PhD thesis, Université Paris-Dauphine., 1998.

[26] S. Masnou and J. Morel. Level-lines based Disocclusion. In *IEEE Intenrational Conference on Image Processing*, pages 259–263, Chicago, IL, 1998.

[27] M. Nitzberg, D. Mumford, and T. Shiota. *Filtering, Segmentation, and Depth*. Springer-Verlag, Berlin, 1993.

[28] L. Wei and M. Levoy. Fast texture synthesis using tree-structured vector quantization. In *Proceedings of SIGGRAPH*, pages 479–488, 2000.

[29] S.-C. Zhu, Y. Wu, and D. Mumford. FRAME: Filters, Random fields And Maximum Entropy— towards a unified theory for texture modeling. *International Journal of Computer Vision*, 27:1–20, 1998.