

The final publication is
available at [https://link.springer.com/
chapter/10.1007%2F978-3-662-45523-4_36](https://link.springer.com/chapter/10.1007%2F978-3-662-45523-4_36)

Evolutionary algorithm for dense pixel matching in presence of distortions

Ana Carolina dos-Santos-Paulino¹, Jean-Christophe Nebel², and Francisco Flórez-Revuelta²

¹ Télécom Physique Strasbourg, Université de Strasbourg
Bld. Sébastien Brant, F 67400, Illkirch-Graffenstaden, France
acdossantos@etu.unistra.fr

² Faculty of Science, Engineering and Computing, Kingston University
Penrhyn Road, KT1 2EE, Kingston upon Thames, United Kingdom
{J.Nebel,F.Florez}@kingston.ac.uk

Abstract. Dense pixel matching is an essential step required by many computer vision applications. While a large body of work has addressed quite successfully the rectified scenario, accurate pixel correspondence between an image and a distorted version remains very challenging. Exploiting an analogy between sequences of genetic material and images, we propose a novel genetics inspired algorithm where image variability is treated as the product of a set of image mutations. As a consequence, correspondence for each scanline of the initial image is formulated as the optimisation of a path in the second image minimising a fitness function penalising mutations. This optimisation is performed by a evolutionary algorithm which, in addition to provide fast convergence, implicitly ensures consistency between successive scanlines. Performance evaluation on locally and globally distorted images validates our bio-inspired approach.

Keywords: Evolutionary algorithm, dense pixel matching, unrectified images, distorted images

1 Introduction

Despite indubitable progress in last decades, success of current image processing algorithms is largely constrained to controlled environments. In addition, attempting to control the huge number of parameters involved in scene variability is a very clumsy and inefficient way of dealing with real-life situations. In contrast, in the field of bioinformatics, dealing with data variability is at the core of most algorithms since genetic mutations are a reality which cannot be ignored. Based on this observation, the authors have worked on a novel genetics-inspired paradigm for image analysis. This new paradigm relies on the idea that by making an analogy between sequences of genetic material and images, image variability can be interpreted as the product of image mutations. Since many computer vision systems rely either on pixel matching or optical flows, this task

has been the core of our investigations [1]. The novel dense pixel matching algorithm we proposed based on this paradigm has demonstrated its robustness not only to camera rotation and translation, but also to local and global distortions [1], so that stereo matching can be freed from the constraint of working with rectified images.

Despite these achievements, the proposed algorithm displays two main limitations: lack of consistency between matches of successive scanlines and a high computational complexity due to the selection of a dynamic programming algorithm to optimise pixel matching. In this work, we propose to address those drawbacks by performing optimisation using an algorithm fitting our bio-inspired paradigm, i.e. a cellular evolutionary algorithm.

A cellular evolutionary algorithm (cGA) is a specific type of evolutionary algorithm where the individuals in the populations are connected establishing a neighbourhood relation between them. Particularly, individuals are conceptually set in a toroidal mesh, and are only allowed to recombine with close individuals [2]. This model fits well with the structure of a image composed by neighbouring lines. Besides, the distortion of a line is quite similar to the distortions of the neighbouring lines. This fact introduces a modification to the original cellular evolutionary model: there is not a single fitness function to optimise, but each individual will optimise the distortion for a specific line in the images.

2 Related work

The main application of dense pixel matching has been 3D reconstruction from a pair of stereo images. A large body of research has been devoted to the scenario which assumes that images have been rectified so that the problem can be reduced to finding correspondences between a pair of scanlines, see reviews [3][4]. Image rectification has been mainly focused on addressing global distortions associated with camera rotation and translation, and lens distortions. Standard approaches include planar [5], cylindrical [6] and spherical [7] rectifications. In addition to their reliance on finding a set of accurate matching points, pixel interpolations and usage of simple camera models, they are not able to deal with local distortions such as those produced by raindrops and dust.

Although the alternative is the design of dense pixel matching approaches which do not require prior image rectification, very few algorithms have been proposed. [8] offered a solution using multi-resolution image correlation. However, since it was developed to address the particular task of 3D reconstruction of a unique convex object, applications have focused on either face or body part modelling [9][10]. [11] offered a variation of a motion estimation algorithm used for JVT/H.264 video coding to perform a 2-dimensional search. However, the lack of contextual constraints makes the matching process particularly difficult in poorly textured regions. More recently, [12] presented a modification of how the cost volume is created during matching which, they claim, can be integrated in any disparity estimation framework. However, since they only present results on slightly misaligned images, behaviour of their approach in more complicated

scenarios remains unknown. Finally, we proposed a novel algorithm, whose fitness function was inspired by our genetics-inspired paradigm. Although it has demonstrated robustness to many camera transformations and distortions [1], it suffers from using a dynamic programming approach for optimisation of the pixel matching process. First, disparity maps display horizontal streaking due to the absence of consistency constraint between successive matched lines. Second, it has a high computational complexity ($O(n^3)$, assuming an image of size n^2).

Although dynamic programming techniques based on tree structures [13][14] have been proposed to allow optimisations across both vertical and horizontal dimensions, they do not provide true 2D optimisation since optimisation is performed along a tree structure instead of a whole image. Moreover, they display a higher computational cost and have only been applied to the rectified scenario. As a consequence, we have investigated alternative approaches to optimise our fitness function. Although dense pixel matching using a genetic algorithm has only been proposed in the rectified image scenario [15][16], an elegant approach allowing optimising matching fitness functions has recently shown a 20% quality improvement while performing fast convergence [17]. In this work, we propose to adopt a similar scheme, but adapted to the unrectified scenario.

3 Bio-inspired algorithm

Data explosion in sequencing of genetic material gives researchers the opportunity to compare sequences of genetic material to establish evolutionary relationship between proteins. Since protein sequences have an average length of 400 characters and mutate through substitution, insertion and deletion of characters, the alignment of a protein pair is not a trivial matter. The 'Needleman-Wunsch' algorithm [18] has provided an effective automatic method to produce an exact solution to the global alignment of two protein sequences which is still at the core of the latest protein search tools. It is based on a dynamic programming (DP) approach which optimises the global alignment of character strings according to a scoring function taking into account possible mutations. In practice, alignments are produced in two steps. First, a 2D scoring matrix is filled where each cell stores the maximum value which can be achieved by extending a previous alignment. This can be done either by aligning the next character of the first sequence with the next character of the second sequence ('match') or extending either sequence by an empty character to record a character insertion or deletion ('gap'). Second, a 'backtracking' process extracts the optimal path in the matrix, which leads to finding the best alignment between the two sequences.

An analogy can be drawn between aligning protein sequences and matching pixels belonging to scanlines, since both tasks aim at establishing optimal correspondence between two strings of characters: the second image of a pair can be seen as a mutated version of the first image where noise, distortions and individual camera sensitivity alter pixel values (i.e. character substitutions); and a different view angle reveals previously occluded data and introduces new occlusions (i.e. insertion and deletion of characters). In earlier work, taking advantage

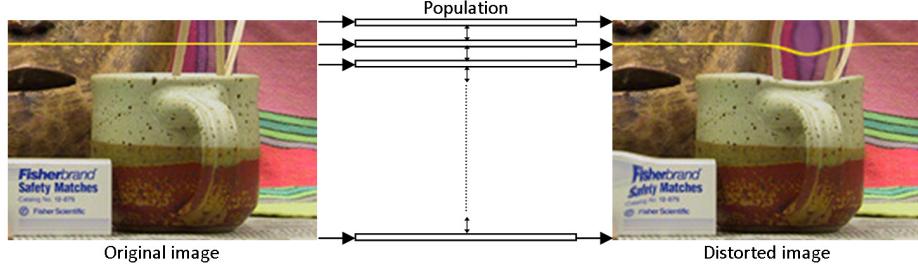


Fig. 1: Each individual in the population represents the path of a scanline in the distorted image. Individuals are linearly connected.

of this analogy, the authors proposed a novel dense pixel matching algorithm able to find correspondence between unrectified images [1]. That approach was shown robust not only to camera rotation and translation, but also to local and global distortions since, instead of restricting itself to finding pixel correspondences between scanlines, it does it between a scanline and an entire image. This was achieved by using a 3D scoring matrix, which allowed taking into account a larger range of 'mutations', see Table 1, so that image distortions could be addressed.

Since pixel matching relies on a DP algorithm operating in a 3D matrix, that approach has a high computational complexity. Moreover, the processing of each scanline independently does not ensure any consistency between matches of neighbouring scanlines. Here, we address those drawbacks by replacing the DP based optimisation by an evolutionary algorithm. Its main principle is the evolution, for each scanline of the first image, of a path within the second image which optimises a scoring function maximising pixel matches and minimising gaps. Moreover, since our evolutionary algorithm relies on recombination between close neighbours, this proposed optimisation implicitly provides consistency between successive scanlines.

4 Evolutionary proposal

Let I_1 and I_2 be a pair of images composed of n lines, where I_2 is a distorted version of I_1 . Each individual in the population represents the path of one line in I_1 (scanline) within I_2 (Fig. 1). Therefore, the number of individuals in the population is equal to the number of lines in the image I_1 .

The approach taken is very similar to the behaviour of a linear Cellular Genetic Algorithm with the exception that fitness varies between individuals. However, as neighbouring scanlines have similar distortions, each individual will optimise a similar fitness function. Similarly to cGAs, recombination will take place between close neighbours propagating good solutions to the neighbourhood (Algorithm 1). As a result, the final population will represent distortions between images.

Algorithm 1 Evolutionary process

```

Initialise the population (see Section 4.3) with a number of individuals equal to the
number of lines in the image
repeat
    Select an individual  $I$  at random
    Select a couple of parents  $I_1$  and  $I_2$  in the neighbourhood of  $I$  (Section 4.4)
    Recombine  $I_1$  and  $I_2$  by crossover generating  $I_{new}$  (Section 4.5)
    Mutate  $I_{new}$  (Section 4.6)
    if  $fitness(I_{new}) < fitness(I)$  then
        Substitute  $I$  by  $I_{new}$ 
    end if
until an ending condition is fulfilled

```

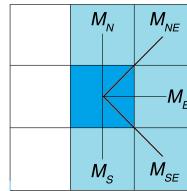


Fig. 2: Different possible matches according to the direction. Stay directions follow the same nomenclature.

Besides, since the individuals correspond to different lines of the image, the topology of the population is linear, unlike standard cGAs operating in 2D which evolve in a toroidal structure. That is coherent, as the first and the last lines of an image are neither spatially related, nor likely to present a similar distortion pattern.

4.1 Individual's representation

Each individual represents the path that a scanline in the original image follows in the distorted image. Each gene takes one of the 11 values described in Table 1. A match is achieved when a pixel in the scanline has a correspondence in the boundaries of the previous pixel in the distorted image. This match can be located in either the north, south, east, northeast or southeast direction of the pixel previously analysed (Fig. 2). Moreover, due to distortion or occlusion (if the images are captured from different view points), a pixel of the original image may not be found in any of the previously defined adjacent positions or a pixel in the distorted image may not have any correspondence in the original image. In the first case, a *stay* (making reference to *stay in the scanline, while moving in the distorted image*) may be placed in the sequence of directions. In the second case, a *gap* is placed in the distorted image.

Gene values	Representation	Penalty	Motion in the original image	Motion in the distorted image
Match North	M_N	$d(P_{Orig}, P_{Dist})$	\rightarrow	\uparrow
Match South	M_S	$d(P_{Orig}, P_{Dist})$	\rightarrow	\downarrow
Match Northeast	M_{NE}	$\sqrt{2} \cdot d(P_{Orig}, P_{Dist})$	\rightarrow	\nearrow
Match Southeast	M_{SE}	$\sqrt{2} \cdot d(P_{Orig}, P_{Dist})$	\rightarrow	\searrow
Match East	M_E	$d(P_{Orig}, P_{Dist})$	\rightarrow	\rightarrow
Stay North	S_N	g	—	\uparrow
Stay South	S_S	g	—	\downarrow
Stay Northeast	S_{NE}	$\sqrt{2} \cdot g$	—	\nearrow
Stay Southeast	S_{SE}	$\sqrt{2} \cdot g$	—	\searrow
Stay East	S_E	g	—	\rightarrow
Gap	G	g	\rightarrow	—

Table 1: Possible gene values, and associated penalties and motions considered in Algorithm 2.

Algorithm 2 Fitness calculation

```

Let  $P_{Orig} = (x_{Orig}, y_{Orig})$  and  $P_{Dist} = (x_{Dist}, y_{Dist})$  be the first pixels in associated lines in both the original and the distorted image. Therefore,  $x_{Orig} = 1$ ,  $x_{Dist} = 1$  and  $y_{Orig} = y_{Dist}$ 
Set fitness=0
for  $i = 1$  to length of the individual do
    Obtain the value  $V_i$  for the gene  $i$  in the individual
    Increase the fitness with the penalty associated to  $V_i$  following Table 1 between  $P_{Orig}$  and  $P_{Dist}$  where  $d(P_{Orig}, P_{Dist}) = \|RGB(P_{Dist}) - RGB(P_{Orig})\|$  and  $RGB(P)$  is the RGB value of  $P$ 
    Update the coordinates for  $P_{Orig}$  and  $P_{Dist}$  according to the motions described in Table 1
end for

```

4.2 Fitness

The fitness function is a measure of the discrepancy between a scanline and its correspondence in the distorted image. It is calculated following Algorithm 2 where successive comparisons are carried out between each pair of pixels that are associated by the path represented by the individual.

We follow the same approach as [1] where matches or stays in NE and SE directions are more penalised since they imply moving by a distance of $\sqrt{2}$ pixels in the image. The value g represents a constant penalty when there is no match.

4.3 Initialisation

Each individual is initialised to represent a trajectory beginning from the left of the distorted image. We propose two types of initialisation strategies: either the selection of a sequence of symbols at random, or the generation of a "locally optimised" path, where the addition of each new symbol relies on keeping the individual fitness minimal.

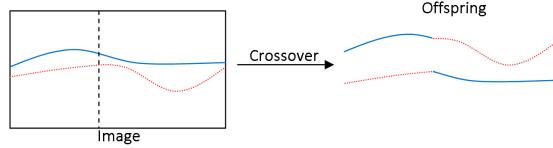


Fig. 3: Illustration of an one-point crossover.

The path creation process is completed when (i) the path reaches one of the borders (top, bottom rows or right column) of the distorted image and (ii) it provides correspondence for all pixels of the scanline. If a border is reached first, gaps are included at random positions in the chromosome until all scanline pixels have correspondence. Otherwise, if there is a correspondence for every pixel in the scanline, stay symbols are included at random in the chromosome until the path reaches one of the borders.

4.4 Selection of parents

As mentioned previously, a new individual I is evolved from a pair of individuals in its neighbourhood³. Selection of the parents could be performed in different ways. For instance:

- according to their proximity to I : the probability of an individual to be chosen is calculated according to a gaussian distribution centered in I , or
- according to their fitness using a roulette-wheel method, where the probability of an individual is a function of its fitness and its neighbours' as described by:

$$P_{I_i} = \frac{\text{fitness}(I_i)}{\sum_{\forall j \in \text{Neighbourhood}(I_i)} \text{fitness}(I_j)} \quad (1)$$

4.5 Crossover

Since individuals have different lengths and representations (paths), which makes their alignments difficult, a typical crossover strategy, where a point (or more) in both parents is selected and the different parts are swapped, is not suitable. Therefore, we propose to select one or more columns in the initial image so that offsprings are generated by swapping the path portions that are defined by those columns (Figure 3). With this procedure, we ensure that each offspring completes a path from the left to the right part of the image.

4.6 Mutation

Mutations produce local changes in the path of an offspring. Those changes can take the following forms:

³ Note that the neighbourhood also includes individual I

- gene alteration by selecting at random an alternative direction,
- gene deletion,
- gene replication over a contiguous interval of the path,
- simultaneous deletion of a gene pair *gaps/stays* in the scanline, as they correspond to inverse operations, and
- local optimisation by substituting an interval of the path by an optimised path following the methodology presented in Section 4.3.

If the mutation process leads to the generation of an individual whose path does not reach the end of the scanline or one of the borders of the distorted image, it is corrected by employing a process similar to the one presented in Section 4.3. Alternatively, if the path exceeds either the length of the scanline or the distorted image, it is cropped.

5 Experimentation

In order to evaluate our evolutionary based dense pixel matching algorithm we have tested it with different image pairs presenting either global distortions, i.e. distortions affecting the whole image, or local distortions where various distortion filters are applied to different areas of the image. Given that pixel values for each colour channel range from 0 to 255 we have set a penalty $g = 181$, similarly to [1], when a 'gap' or a 'stay' is included in a path. We have also considered all the mutation types stated in Section 4.6 with equal probability. As stopping condition, the evolutionary process finishes if there is no changes in the population for 10,000 generations (i.e. the creation of 10,000 new individuals).

Figure 4 presents matching results between an image 4a and its global sinusoidal distortion 4b. While Figure 4c shows paths representing different individuals in the image, Figure 4d displays how the individuals represent the distortion between the original and the distorted images. All individuals, except those in the top and bottom lines, converged towards very similar distortions. In those border areas, there is no continuity in the distorted stripes. Therefore the algorithm minimises the fitness function by either inserting gaps or jumping to a neighbouring stripe with a similar colour. Finally, Figure 4e shows a reconstruction of the original image using the distorted image and the paths coded in the individuals. The original image is recovered quite satisfactorily, except in the border areas previously mentioned.

Figure 5 shows results for an image pair, proposed by [1], where local distortions were applied on the original image (Figure 5c). Figure 5e and 5f highlight the algorithm's ability to detect and correct the distortion on the rim of the mug. However, in this case, it fails to address the distortion affecting the three straws. One can speculate that, due to the choice of fitness function, the algorithm finds easier to deal with distortions involving matches than gaps or stays.

Next, we studied how the choice of initialisation strategy, see section 4.3, affects performance. As seen in Figure 6, usage of locally optimised initialisation instead of random one conducts to the generation of an original population which is composed of paths representing quite well the actual distortion. One

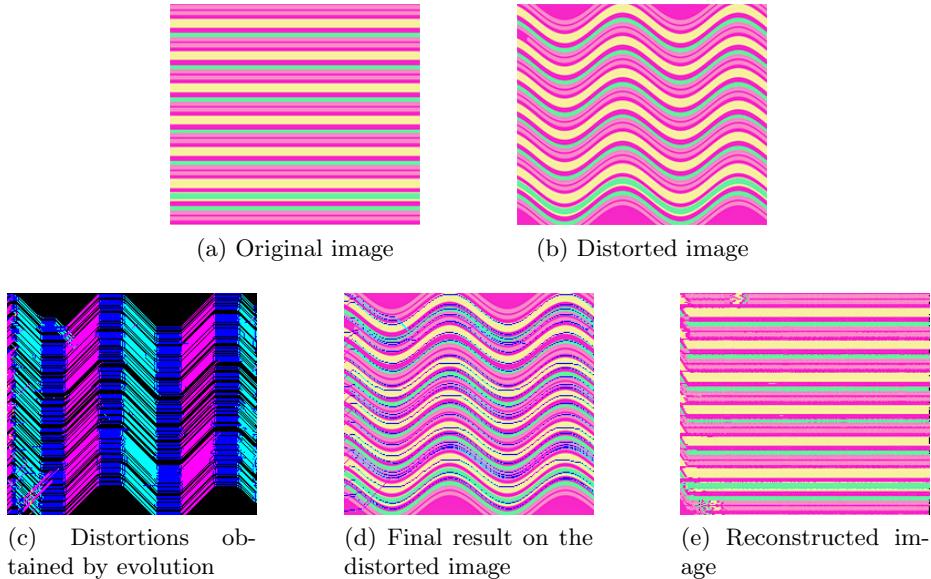


Fig. 4: Image matching in presence of a global sinusoidal distortion. In (c) colours represent different directions. The reconstructed image (d) is generated by pasting the distorted image pixels according to the path estimated for each line.

may wonder if this optimised initialisation compromises exploration of the solution space and leads to convergence towards a worse fitness. Results in Table 2 reveal that actually optimised initialisation leads to better solutions and faster convergence of the algorithm.

Another important aspect of the algorithm is parents selection. First, Figure 7 shows how the fitness function evolves according to the size of the neighbourhood considered when selecting parents by roulette taking into account the fitness of the individuals. In the case of the stripes image (Figure 7a) where the distortion is global, the larger the neighbourhood the better the results, as good solutions propagate faster. In the case of local distortions (Cones image - 7b), the best result is obtained for a neighbourhood of size 15, whereas there is no significant difference between usage of sizes 5 and 25. The best neighbourhood seems to be related to the actual size of the distorted areas. These figures also confirm the superiority of locally optimised initialisation over the random one, since in both cases it provides faster convergence towards a lower fitness value. Note that results for random initialisation of all individuals are not shown for the Cones images as they would not fit on the graph.

Second, comparison was performed between two of the methods proposed in Section 4.4 for the selection of parents, i.e. roulette and proximity to the individual to be evolved. As shown on Figure 8, no significant differences are observed. Therefore, the evolution seems to be more affected by the size of the neighbourhood than the method used to select the parents.

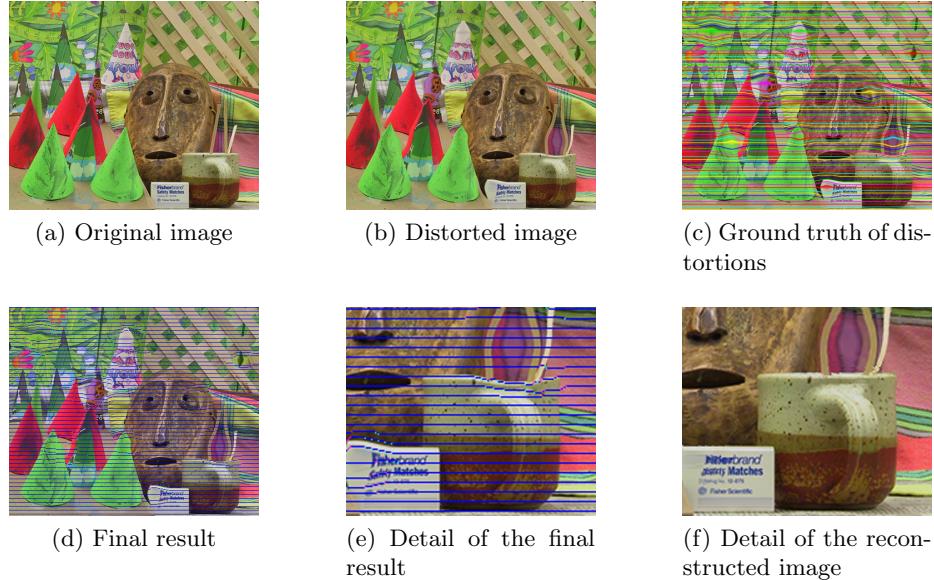


Fig. 5: Results obtained with an image with local distortions.

6 Conclusion

We have introduced a novel dense pixel matching algorithm suitable for the unrectified scenario. Based on a bio-inspired approach, our main contribution has been the design of an evolutionary algorithm able to optimise a different fitness function for each scanline while ensuring consistency between neighbouring lines. As demonstrated in experiments involving locally and globally distorted images, the proposed approach is valid, since processing converges towards satisfactory solutions which do not display any horizontal streaking. Moreover, study of different population initialisation and parents selection strategies has revealed that locally optimised initialisation provides a better performance, while the selection of parents is more affected by the choice of the chromosome pool than the method used to extract the actual genitors. As future work, we propose to further develop our system so that it could produce 3D reconstruction in real time from data generated by two uncalibrated video cameras.

References

1. Thevenon, J., Martinez del Rincon, J., Dieny, R., Nebel, J.C.: Dense pixel matching between unrectified and distorted images using dynamic programming. In: Intl. Conference on Computer Vision Theory and Applications. (2012) 216–224
2. Alba, E., Dorronsoro, B.: Cellular genetic algorithms. Volume 42. Springer (2008)

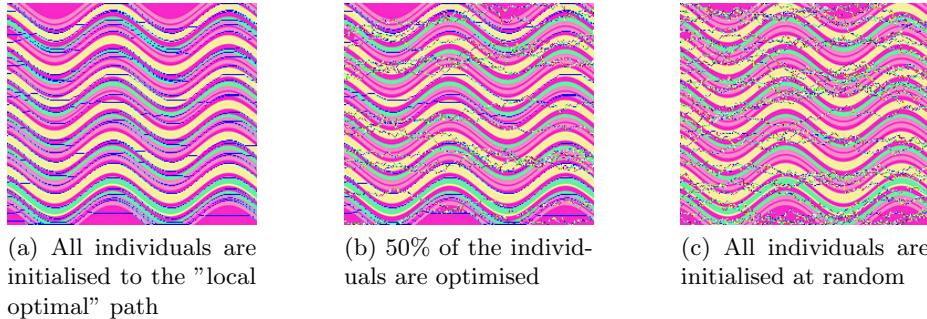
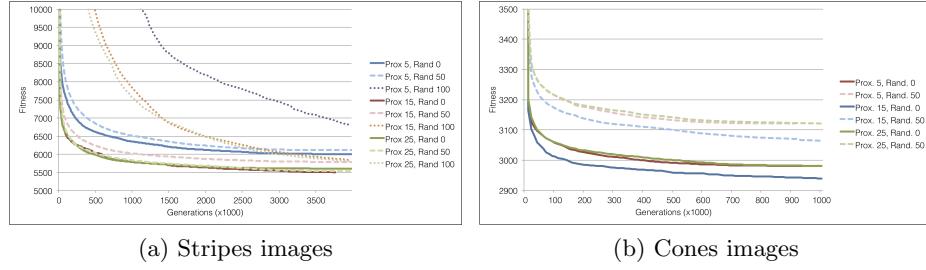


Fig. 6: Initial populations generated from different initialisation methods

Table 2: Comparison between different initialisation methods. Ten runs were carried out with the stripes images using the roulette method with a neighbourhood size 5 to select parents in the crossover.

Percentage of random individuals created	Iterations until convergence (in millions)			Initial fitness	Final fitness		
	Average	Best	σ		Average	Best	σ
0	2.91	2.00	0.74	12,739.58	6,003.22	5,802.13	107.63
50	3.06	2.24	0.54	48,818.54	6,113.75	5,992.04	98.75
100	7.92	5.07	2.23	58,397.66	6,720.27	5,983.65	739.10

3. Scharstein, D., Szeliski, R.: A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. *Intl. Journal of Computer Vision* **47**(1-3) (2002) 7–42
4. Tippett, B., Lee, D., Lillywhite, K., Archibald, J.: Review of stereo vision algorithms and their suitability for resource-limited systems. *Journal of Real-Time Image Processing* (2013) 1–21
5. Hartley, R.I.: Theory and practice of projective rectification. *International Journal of Computer Vision* **35**(2) (1999) 115–127
6. Roy, S., Meunier, J., Cox, I.J.: Cylindrical rectification to minimize epipolar distortion. In: *Proceedings of the 1997 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. (1997) 393–399
7. Wan, D., Zhou, J.: Self-calibration of spherical rectification for a ptz-stereo system. *Image and Vision Computing* **28**(3) (2010) 367 – 375
8. Zhengping, J.: On the multi-scale iconic representation for low-level computer vision systems. PhD thesis, The Turing Institute and the U. of Strathclyde (1988)
9. Khambay, B., Nebel, J.C., Bowman, J., Ayoub, A., Walker, F., Donald, Hadley, D.: A pilot study: 3D stereo photogrammetric image superimposition on to 3D CT scan images - the future of orthognathic surgery. *The International Journal of Adult Orthodontics and Orthognathic Surgery* **17**(4) (2002) 331–341
10. Cockshott, W.P., Hoff, S., Nebel, J.C.: Experimental 3-D digital TV studio. *Vision, Image and Signal Processing, IEE Proceedings -* **150**(1) (2003) 28–33
11. Nalpantidis, L., Amanatiadis, A., Sirakoulis, G., Kyriakoulis, N., Gasteratos, A.: Dense disparity estimation using a hierarchical matching technique from uncalibrated images. *Journal of Visual Communication and Image Representation* **14**(2) (2003) 130–142



(a) Stripes images

(b) Cones images

Fig. 7: Influence of parents selection and initialisation strategies (*Prox* indicates the number of individuals considered in the neighbourhood and *Rand* the percentage of individuals initialised at random)

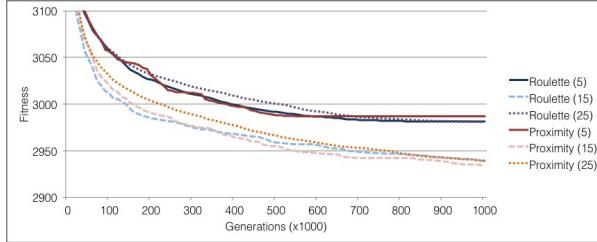


Fig. 8: Influence of the selection method and the neighbourhood size

- brated stereo vision. In: IEEE International Workshop on Imaging Systems and Techniques. (2009) 427–431
12. Rzeszutek, R., Tian, D., Vetro, A.: Disparity estimation of misaligned images in a scanline optimization framework. In: IEEE International Conference on Acoustics, Speech and Signal Processing. (2013) 1523–1527
 13. Veksler, O.: Stereo correspondence by dynamic programming on a tree. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Volume 2. (2005) 384–390
 14. Deng, Y., Lin, X.: A fast line segment based dense stereo algorithm using tree dynamic programming. In: Computer Vision ECCV 2006. Volume 3953 of Lecture Notes in Computer Science. Springer Berlin Heidelberg (2006) 201–212
 15. Saito, H., Mori, M.: Application of genetic algorithms to stereo matching of images. Pattern Recognition Letters **16**(8) (1995) 815 – 821
 16. Han, K.P., Song, K.W., Chung, E.Y., Cho, S.J., Ha, Y.H.: Stereo matching using genetic algorithm with adaptive chromosomes. Pattern Recognition **34**(9) (2001) 1729 – 1740
 17. Kiperwasser, E., David, O., Netanyahu, N.S.: A hybrid genetic approach for stereo matching. In: Proceeding of the 15th Genetic and Evolutionary Computation Conference, New York, USA, ACM (2013) 1325–1332
 18. Needleman, S.B., Wunsch, C.D.: A general method applicable to the search for similarities in the amino acid sequence of two proteins. Journal of Molecular Biology **48**(3) (1970) 443 – 453