Coevolutionary strategies in area-based stereo

John Y. Goulermas[a], Panos Liatsis[b]

[a]Knowledge Support Systems Limited, Peter House, Oxford Street, Manchester M1 5AN, United Kingdom

[b]Intelligent Systems and Sensing Laboratory, Dept of Electrical Engineering & Electronics, UMIST, PO Box 88, Manchester M60 1QD, United Kingdom
yannisg@kssg.com, panos.liatsis@umist.ac.uk

Abstract: *A novel system for depth estimation is proposed with the use of Symbiotic Genetic Algorithms for the continuous problem of disparity surface approximation. The approach is based on the decomposition of the entire surface to very small non-overlapping patches described by low-order bivariate polynomials and the use of symbiotic optimisation to enforce smoothness at the boundaries of these patches, so that the entire surface can be approximated in a smooth piecewise fashion by functionals of local support. Such optimisation is amenable to a massive parallel implementation, since each patch is optimised by a different execution unit and each unit communicates through its cost function only with its four-connected neighbours. The method makes use of various existing crossover and mutation schemes for real-valued chromosome representations and a new problem-specific mechanism for generating and hybridising the initial populations. The proposed multi-objective cost function enforces photometric similarity and smoothness between the patch boundaries at a local scale, which in the long term give rise to a globally smooth disparity surface.*
Keywords: *Area-based stereo, co-evolutionary optimization, disparity surface approximation.*

1. INTRODUCTION

This paper proposes a novel versatile version of the SGAs for piecewise surface approximation in area-based stereo. The algorithm decomposes the entire disparity surface into a large number of surface patches and, by allocating a separate population to optimise each patch, it employs symbiotic interactions to enforce smoothness and reduce mismatches at a global scale. Section 2 describes the preprocessing used to reduce the photometric discrepancies between the images of the stereo-pair. Section 3 describes the proposed SGAs and specifically, how each patch is represented by a real-valued vector encoding the coefficients of a low-order polynomial, and also, the composite fitness function comprised by self and symbiont contributions. Finally, Section 4 evaluates the proposed method and Section 5 provides the conclusions and suggestions for further work.

2. PREPROCESSING

It is possible for the two parts of the stereo pair to have different photometric properties. This could complicate the comparison of intensity measurements. To account for such discrepancies, we use a simple yet effective technique[2]. The photometric variation is modelled through a constant additive and a multiplicative factor α and β, respectively, via:

$$L(y,x) = \alpha \cdot R(y, x - d(y,x)) + \beta \qquad (1)$$

where (y,x) is a pixel of the reference image L and d(y,x) its disparity. Assuming low occlusions and noise, the intensity histograms can be used to correct the likely variations. First, the histograms for both left and right image are constructed. Then, the n percentile points of L are plotted against the same percentile points of R. These are the intensities that correspond to the first $w \cdot h/n$, $2 \cdot w \cdot h/n$, …, $w \cdot h$ pixels of L and R. Subsequently, a piecewise linear approximation is performed between these points.

3. SGAS FOR AREA-BASED STEREO

The concept of symbiosis can be used to decompose the problem of finding the complete dense disparity map to a number of subproblems (species), assigning each subproblem to a different GA population and then simulating evolution using fitness functions which take into account the interdependencies between the defined subproblems. One way to achieve this is to partition the reference image into rectangular patches, each of size $w_p \times h_p$ and try to locate the disparities within each patch. In this case, each chromosome can be a 1D or a 2D structure of $w_p \cdot h_p$ genes encoding disparities for all pixels within the patch it corresponds, thus giving a search space of approximate size $(d_{max}-d_{min}+1)^{wp \cdot hp}$.

The proposed algorithm encodes continuous disparity functions, as opposed to discrete disparity values and also uses the interdependencies between the patches in order to enforce smoothness not only within each patch, but also at the patch boundaries in order to reduce the graininess effect. Furthermore, explicit within-patch smoothness evaluation is avoided by modelling the local disparities using 2D polynomial surfaces of low-order.

3.1. Genome Encoding

The $w \times h$ left retina L is partitioned into adjacent non-overlapping rectangular patches each of size $w_p \times h_p$, shown in Fig. 1, so that each patch is centred at the point (j,i) of a uniformly spaced grid. We use $d_{ji}(y,x)$ to denote a 2D continuous function with domain limited within the boundaries of the patch it represents; we refer to this at the disparity functional. In this way, the range of each $d_{ji}(y,x)$ corresponds to the set of disparities at L-coordinates, with the domain $[y_j^-,y_j^+] \times [x_i^-,x_i^+] \equiv [h_p \cdot j, h_p \cdot (j+1)-1] \times [w_p \cdot i, w_p \cdot (i+1)-1]$. Note, that image points (y,x) are indexed by $y=0,\ldots,h-1$ and $x=0,\ldots,w-1$, while grid-points (j,i) are indexed by $j=0,\ldots,h_g-1$ and $i=0,\ldots,w_g-1$, where $h_g=\lfloor h/h_p \rfloor$ and $w_g=\lfloor w/w_p \rfloor$.

At this point, the entire disparity surface has been decomposed into a set of much smaller disparity surfaces, each represented by a chromosome structure. As mentioned above, instead of

encoding each chromosome as a 1D or 2D string of discrete disparity values with $w_p \cdot h_p$ genes, we make use of bivariate polynomials of degree N to continuously model each disparity functional as:

$$d_{ji}(y,x) = \sum_{a+b \le N} z_{ab} \cdot y^a \cdot x^b \qquad (2)$$

where z_{ab} is the coefficient of each term. Note, that it can be shown[3] that the number of positive integer solutions of the equation a+b≤N is given by $\tau(N) \equiv \binom{N+2}{N} = \frac{(N+1)(N+2)}{2}$; this number corresponds to the number of coefficients in an N_{th} degree bivariate polynomial. Thus, we are allowed to model each chromosome as the concatenation of these coefficients, so that each chromosome is simply a vector of real numbers; e.g. $[z_{20},z_{02},z_{11},z_{10},z_{01},z_{00}]$ of length $\tau(2)=6$.

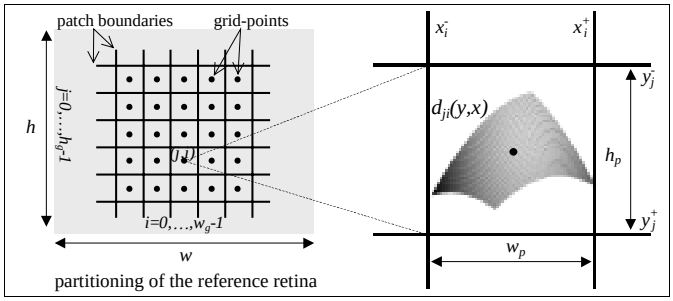

Fig. 1: Partitioning of L into $w_g \times h_g$ rectangular patches. The disparities within each patch are computed by the values of its disparity functional $d_{ji}(y,x)$ in the domain $[y_j^-,y_j^+] \times [x_i^-,x_i^+]$.

3.2. The Genetic Operators

The g_{child} new offspring at each population are created by either crossover or mutation, independently. At each generation t, each GA performs a number T of cycles of selecting the next genetic operator to be applied until all g_{child} offspring are collected. At each cycle, crossover is applied with probability p_c and mutation with probability $(1-p_c)$. Note, that crossover generates an average of 1.75 offspring, while mutation 1 child per application. Since, $p_c \cdot T \cdot 1.75 + (1-p_c) \cdot T \cdot 1$ must be equal to g_{child}, we need $T=g_{child}/(0.75 \cdot p_c+1)$ cycles in total to collect the g_{child} offspring. Thus, crossover and mutation produce $175 \cdot p_c/(0.75 \cdot p_c+1)\%$ and $(1-p_c) \cdot 100/(0.75 \cdot p_c+1)\%$ offspring, respectively. We use a value of $p_c=0.6$, which gives 72.4% and 27.6% for the two operators, respectively. Since, at the initial generations adequate diversity exists, the exploitation ability of crossover is more required. Later on, when diversity is decreased the exploration of new regions by the mutation operator is more useful. Thus, we have used a heuristic to take into account this observation, by halving p_c at a predefined generation threshold $t=t_c$, (typically set to 100). For the generations $t_c,\ldots,t_{max}$ a value of $p_c=0.3$ changes the contributions from crossover and mutation to 42.9% and 57.1% of g_{child}, respectively.

3.3. Symbiotic Credit Assignment

Each of the $w_g \cdot h_g$ symbiotic populations try to locate, for each individual patch, the set of coefficients which produce the optimal disparity surface within the patch. A first objective is to derive surfaces which correspond to matches between pixels that exhibit high photometric similarity. Given a set of coefficients, each disparity surface $d_{ji}(y,x)$ produces a discrete set of individual matches between all left pixels $(y,x) \in [y_j^-, y_j^+] \times [x_i^-, x_i^+]$ and the right pixels $(y, l-d_{ji}(y,x))$. The overall error for all the matches corresponding to the disparity surface within each patch is defined as a summation of the photometric dissimilarities between the participating matches. A large number of metrics, such as the ones discussed in Section 2, can be used. The simplest one is the S_{SD} which measures the squared intensity difference between the pixels. Although this metric is the weakest to used, since as a pixel-to-pixel comparator is prone to photometric discrepancies, noise and occlusions, many methods[6] use it due to its simplicity and amenability for incorporation to a cost function subjected to an energy minimisation framework. Since GAs do not require differentiability of the performance function such as the gradient and variational calculus based methods, any other correlation metric (e.g., S_{NSSD}, S_{NMSD}, S_{ZNCC}) can be used. Nevertheless, the increased robustness of these metrics incurs a proportionate increase of the computational overhead required by the multiple evaluations of the fitness functions.

As the first objective of the cost function governing the optimisation of each patch centred at grid-point (j,i), we define the function $f_M(c_{jik})$, which returns the average S_{SD} value of all the matches within the patch, generated by the $d_{ji}(y,x)$ disparity surface defined by the $\tau(N)$ coefficients of each k_{th} chromosome c_{jik} of population P_{ji}. For notational simplicity, we use $c_{jik}(y,x)$ to denote the disparity value of this chromosome at pixel (y,x). However, certain matches may produce out-of-range disparities. There is also a chance that certain disparity values may have left pixels matched with right pixels out of the boundaries of the right retina. We choose to penalise the former case by a constant cost per pixel ξ_1, since it violates the problem constraints. For the latter case, we also allocate a high fixed cost ξ_2, since it represents non-measurable elements and only pixels from the left and right boundaries of the left and right retinas, respectively, may lie outside the field of view. Finally, the first objective f_M is formulated as:

$$f_M(c_{jik}) = \frac{1}{w_p \cdot h_p} \cdot \sum_{y=y_j^-}^{y_j^+} \sum_{x=x_i^-}^{x_i^+} \begin{cases} \xi_1 & \text{if } c_{jik}(y,x) \notin [d_{min}, d_{max}] \\ \xi_2 & \text{if } x - c_{jik}(y,x) \notin [0, w-1] \\ \left[L(y,x) - R(y, x - c_{jik}(y,x)) \right]^2 & \text{otherwise} \end{cases} \quad (3)$$

The tremendous benefit gained by modelling the entire disparity map as a set of disparity functionals with local support is that low degree polynomials naturally produce smooth local surfaces. Thus, unlike other area-based methods, terms which enforce smoothness within the patches are no more needed. Nevertheless, such a distributed approach, especially with a low-quality metric such as the S_{SD}, can lead to certain problems for two reasons. Firstly, since patches have local support, their size may not be adequate to capture enough information and properly establish photometric similarity. Increasing the size of the patches alleviates this problem, however causes loss of fine elevation detail. Secondly, since the recovery of the depth map is an ill-posed problem, many surfaces may exist giving alike dissimilarity scores. These can cause notable matching errors, not only within the patches, but also at their boundaries, thus giving rise to a mosaic effect in the final disparity map. Although the functionals are smooth within their

boundaries, the transition from one patch to one of its neighbours may exhibit steep disparity jumps and discontinuities.

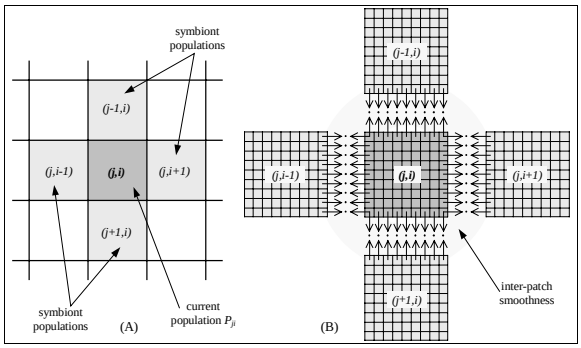

Fig. 2: (A): Topology of the symbiotic dependency between each population P_{ji} and its symbionts $P_{j\pm1\ i\pm1}$. (B): Regularising forces at the patch boundaries which enforce smoothness between the participating polynomial surfaces.

To resolve such effects we regularise the problem by enforcing symbiotic optimisation to proscribe each patch from evolving independently to its neighbours. Inter-patch regularisation is explicitly enforced by adding extra terms to the objective function. Fig. 2.A shows the dependency between each population P_{ji} and its four-connected symbiont populations $P_{j\pm1\ i\pm1}$. To improve smoothness, each surface is driven to obtain boundary disparities similar to those of its adjacent neighbours, to all four sides of the patch; namely, top ($y=y_j^-$), bottom ($y=y_j^+$), left ($x=x_i^-$) and right ($x=x_i^+$), as shown in Fig. 2.B. This is dictated by the term f_{S1} defined as:

$$f_{S_1}(c_{jik}) = \frac{1}{2(w_p+h_p)} \cdot \left\{ \begin{array}{l} \underbrace{\sum_{y=y_j^-}^{y_j^+} \left[c_{jik}(y, x_i^-) - c^*_{j\,i-1}(y, x_{i-1}^+) \right]^2}_{\text{left adjacency}} + \underbrace{\sum_{y=y_j^-}^{y_j^+} \left[c_{jik}(y, x_i^+) - c^*_{j\,i+1}(y, x_{i+1}^-) \right]^2}_{\text{right adjacency}} + \\ \underbrace{\sum_{x=x_i^-}^{x_i^+} \left[c_{jik}(y_j^-, x) - c^*_{j-1\,i}(y_{j-1}^+, x) \right]^2}_{\text{top adjacency}} + \underbrace{\sum_{x=x_i^-}^{x_i^+} \left[c_{jik}(y_j^+, x) - c^*_{j+1\,i}(y_{j+1}^-, x) \right]^2}_{\text{bottom adjacency}} \end{array} \right\} \quad (4)$$

which measures the average of the first-order squared differences in the four sides of the polynomial c_{jik}.

Note, that c^*_{ji} is the currently best chromosome of P_{ji}. In this way, the symbiotic optimisation is sustained by taking into account the interdependencies of each chromosome with the elite unit of its neighbouring species. The benefit gained by f_{S1} is twofold; not only patch boundary discontinuities are reduced, but also each surface itself is regularised to the matches that are compatible with its neighbours.

However, there is another type of irregularity found to be frequent at the patch boundaries. Although the adjacent sides of two surfaces may have equal disparities, they may not be attracted consistently to form a smooth joined surface; this causes a peak effect in the patch boundaries. To eliminate such localised disparity peaks, we propose an additional regularisation term, f_{S2}, based on second-order differential information:

$$f_{S_2}(c_{jik})=\frac{1}{2(w_p+h_p)}\left\{\begin{array}{l}\underbrace{\sum_{y=y_j^-}^{y_j^+}\left\{\left[c_{jik}(y,x_i^-+1)-c_{jik}(y,x_i^-)\right]-\left[c^*_{j\,i-1}(y,x_{i-1}^+)-c^*_{j\,i-1}(y,x_{i-1}^+-1)\right]\right\}^2}_{\text{left adjacency}}+\\ \underbrace{\sum_{y=y_j^-}^{y_j^+}\left\{\left[c^*_{j\,i+1}(y,x_{i+1}^-+1)-c^*_{j\,i+1}(y,x_{i+1}^-)\right]-\left[c_{jik}(y,x_i^++1)-c_{jik}(y,x_i^+)\right]\right\}^2}_{\text{right adjacency}}+\\ \underbrace{\sum_{x=x_i^-}^{x_i^+}\left\{\left[c_{jik}(y_j^-+1,x)-c_{jik}(y,x_i^+)\right]-\left[c^*_{j-1\,i}(y_{j-1}^+,x)-c^*_{j-1\,i}(y_{j-1}^+-1,x)\right]\right\}^2}_{\text{top adjacency}}+\\ \underbrace{\sum_{x=x_i^-}^{x_i^+}\left\{\left[c^*_{j+1\,i}(y_{j+1}^-+1,x)-c^*_{j+1\,i}(y_{j+1}^-,x)\right]-\left[c_{jik}(y_j^+,x)-c_{jik}(y_j^+-1,x)\right]\right\}^2}_{\text{bottom adjacency}}\end{array}\right\} \quad (5)$$

This returns the average of the sum of the individual second-order squared differences in the four sides.

A final observation concerning the above modelling, is that each left point is mapped to a single right point. However, the Uniqueness Constraint (UC) is not guaranteed as a right point can be the conjugate of more than one left points. A further improvement of the quality of the disparity map can be obtained using the Ordering Constraint (OC) to constrain the possible surface shapes. Both constraints can be satisfied simultaneously as follows. Suppose that at a left point (y,x) we have disparity d(y,x) and at the point (y,x+Δx), where Δx is small positive quantity, the disparity is d(y,x+Δx). Then the OC implies that, since x+Δx succeeds x, their conjugate points, with coordinates (y,x-d(x)) and (y,(x+Δx)-d(x+Δx)), respectively, preserve this order. This means that:

$$(x+\Delta x)-d(x+\Delta x)>x-d(x) \;\Rightarrow\; \frac{d(x+\Delta x)-d(x)}{\Delta x}<1 \quad (6)$$

which, for infinitesimal $\Delta x\to 0$, gives the following constraint on the disparity function:

$$\frac{\partial d(y,x)}{\partial x}-1<0 \quad (7)$$

We propose incorporation of the violations of Eq.(7) explicitly in the cost function in the following way. For a bivariate polynomial of degree N, this partial can be easily calculated analytically via:

$$\frac{\partial d_{ji}(y,x)}{\partial x}=\sum_{\left\{\substack{a+b\le N\\ b\neq 0}\right\}} b z_{ab}\cdot y^a\cdot x^{b-1}=\left\{\begin{array}{ll} z_{01} & ,\text{for } N=1\\ 2\cdot z_{02}\cdot x+z_{11}\cdot y+z_{01} & ,\text{for } N=2\\ 3\cdot z_{03}\cdot x^2+2\cdot z_{12}\cdot y\cdot x+z_{21}\cdot y^2+2\cdot z_{02}\cdot x+z_{11}\cdot y+z_{01} & ,\text{for } N=3\end{array}\right. \quad (8)$$

Then, the overall quantification of the violations of Eq.(7) within a given chromosome/surface c_{jik} is measured by the term f_O defined as:

$$f_O(c_{jik})=\frac{1}{w_p\cdot h_p}\cdot\sum_{y=y_j^-}^{y_j^+}\sum_{x=x_i^-}^{x_i^+} f_{OC}\left(\frac{\partial d_{ji}(y,x)}{\partial x}\right) \quad (9)$$

which gives the average of the violations of the UC and OC over all the discrete pixels in the patch. Each violation is measured individually by the thresholding term:

$$f_{OC}(q)=\begin{cases}0 & \text{if } q<1\\ 1+log(q) & \text{otherwise}\end{cases} \quad (10)$$

When Eq.(7) is violated, the nonlinear term in Eq.(10) gives a variable penalty which depends on the magnitude of the violation starting at an offset of 1.

The above four terms of Eqs.(4,5,9,10) are combined in a single weighted sum and compose the following multi-objective function:

$$f(c_{jik})=\underbrace{f_M(c_{jik})+\lambda_O\cdot f_O(c_{jik})}_{\text{self contributions}}+\underbrace{\lambda_{S_1}\cdot f_{S_1}(c_{jik})+\lambda_{S_2}\cdot f_{S_2}(c_{jik})}_{\text{symbiotic contributions}} \quad (11)$$

defined for every chromosome c_{jik} of every population P_{ji}, with λ_O, λ_{S1} and λ_{S2} being fixed user-defined weights. This objective function drives the minimisation for each of the $w_g \cdot h_g$ populations of the collective.

4. RESULTS

Here, we employ the stereo pair of the Pentagon for experimentation with the proposed algorithm. Fig. 3.L shows the w×h=512×512 (left) reference image, partitioned into $w_g \cdot h_g$=46·46=2,116 patches, each of size $w_p \times h_p$=11×11 pixels. To perform the hybridisation scheme, the BLS method is applied using the S_{ZNCC} metric and a correlation window of size 5×5. The resulting disparity map produced by the BLS is shown in Fig. 3.M, where the disparity range $[d_{min},d_{max}]$=[-9,+9] is linearly mapped to the intensity range [255,50]. It is apparent that, together with other mismatches, a large number of pixels have wrongly undefined disparities; these are plotted in 0 (black) intensity. The crude disparity map of Fig. 3.R shows, for each patch (j,i), the constant disparity $\bar{d}_{ji}$ taken to be the median of the within-patch disparity distributions of Fig. 3.M. These are used to hybridise the populations of the initial collective.

Each P_{ji} is evolved for t_{max}=400 generations with g_{size}=70 members of which g_{child}=25 ones are replenished at each generation. Linear ranking uses an upper bound of η_{max}=1.6 and the probability of crossover is set to p_c=0.6; this halves at generation t_c=100. The constant penalties of Eq.(4) are set to $\xi_1=\xi_2$=200. In the first experiment, we make use of bivariate (N=2) polynomials giving $\tau(2)$=6 real-valued genes/coefficients per chromosome. To show the advantage gained from the multi-objective optimisation we present results by omitting certain terms from Eq.(11). Fig. 4.TL shows the final disparity map for generation t_{max} (by linearly mapping $[d_{min},d_{max}]$ to the intensities [255,0]) with weights $\lambda_O=\lambda_{S1}=\lambda_{S2}$=0, that is using only hybridisation and the photometric stability term f_M. The mosaic effect is apparent, as although the disparity surface is smooth within each patch, the inter-patch transitions are mostly discontinuous. Furthermore, many pixels are mismatched since the S_{SD} metric is not adequately robust. Fig. 4.TR shows the final result after including the first-order smoothness term f_{S1} with a weight of λ_{S1}=150. The great improvement is apparent since the disparity transitions are much smoother. However, the boundaries of the patches are still discernible due to the peak effect discussed in Section 4.4. The result from additionally using the second-order smoothness term f_{S2} with λ_{S2}=150, is shown in Fig. 4.BL to have smoothed out the peaks and corrected many

remaining mismatches. Finally, Fig. 4.BR shows the full version of Eq.(11), using in addition the OC penalty term f_O with λ_O=100.

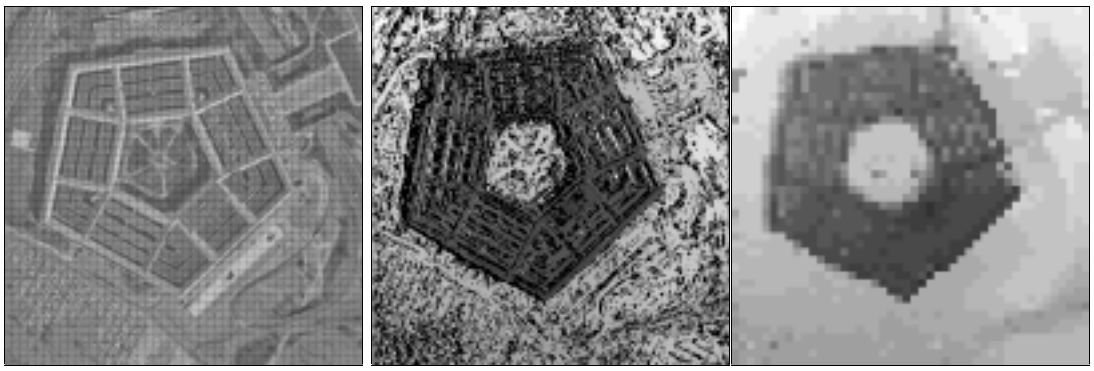

Fig. 3: (L): Decomposition of the left image into $w_g \cdot h_g$ patches. (M): Disparities located by the BLS. (R): Median disparities corresponding to each patch used for hybridisation.

5. CONCLUSION

This paper presented a new algorithm for piecewise surface approximation of the disparity map using low-order bivariate polynomials. Section 2 presented a histogram normalisation method for reducing the photometric discrepancies between the stereo images. Section 3 described the proposed area-based SGA method, including its real-valued chromosome representation scheme, and the multi-objective symbiotic cost function. Results were given in Section 4. The entire optimal surface, the disparity surface is modelled as a piecewise approximation using low-order functionals with small finite support. While self terms in the cost function are concerned with the optimisation of local matching elements, such as the photometric similarity and local surface shape constraints, the symbiotic terms enforce smoothness at the boundaries of the patches, so that transitions from one patch to its neighbouring ones become smooth with reduced discontinuities. In this way, the composite disparity surface is globally smooth and maintains its capability to model accurately local disparity variations. Experiments in Section 4 showed that the system was capable for robust matching and surface approximation. Bivariate polynomials were adequate for accurate modelling. The interaction between the self and symbiotic terms of the cost function was successful and the objectives for good quality matching, smoothness and suppression of discontinuities in the composite disparity maps was satisfied consistently for all participating populations/subproblems in the retina space.

Further work will evaluate the use of alternative disparity surface decomposition schemes for symbiotic continuous optimisation. Each subproblem can be an epipolar slice of the entire disparity map, modelled as a 1D signal spanning the retina width. Basis functions can then be used to approximate each 1D disparity curve. Each separate population can be searching for the optimal weights controlling the shape of its own curve, but symbiotic terms can enforce smoothness across adjacent slices, thus giving rise to a globally smooth and consistent disparity surface.

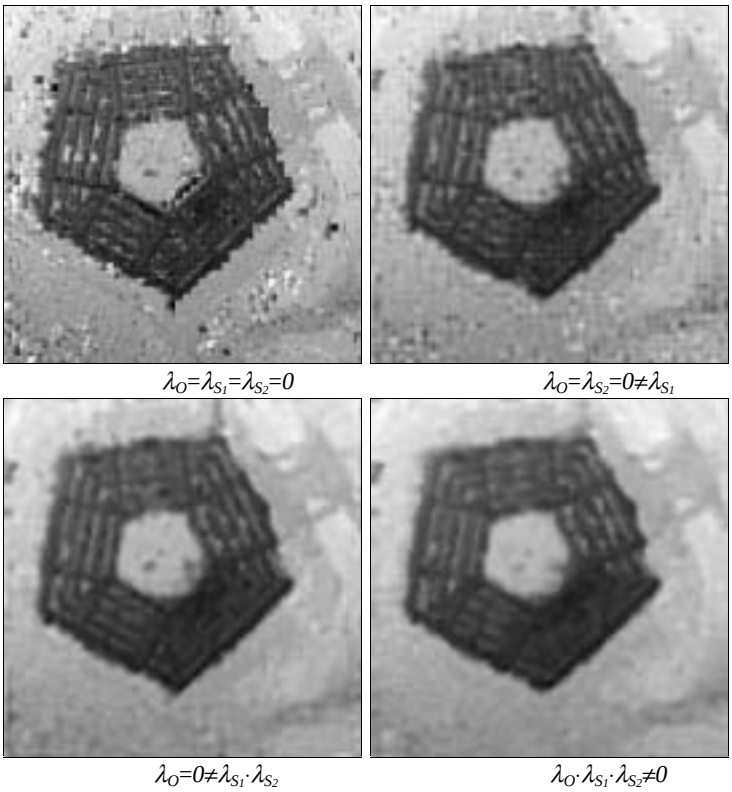

Fig. 4: Linearly interpolated disparity maps for: (TL): using the photometric similarity term only, (TR): using first-order smoothness, (BL): including both first- and second-order smoothness and (BR): full version of Eq.(18) with all terms being used.

REFERENCES

[1] Y. Shirai, *Three-dimensional Computer Vision*, Springer-Verlag, 1987.

[2] I. J. Cox, S.L. Hirogani, S.B. Rao, and B. M. Maggs, "A maximum likelihood stereo algorithm", *Computer Vision and Image Understanding*, **63**, No. 3, pp. 542-567, 1996.

[3] R. M. Haralick, "Digital step edges from zero-crossings of second directional derivatives", *IEEE Trans. PAMI*, **6**, No. 1, pp. 58-68, 1984.

[4] Z. Michalewicz, *Genetic Algorithms+data structures=evolution programs*, Springer-Verlag, 1996.

[5] H. Muhlenbein and D. Schlierkamp-Voosen, "Predictive models for the breeder genetic algorithm. I: Continuous parameter optimization", *Evolutionary Computation*, **1**, No. 1, pp. 25-49, 1993.

[6] B. K. P. Horn, *Robot Vision*, MIT Press, 1986.