České vysoké učení technické v Praze
Fakulta elektrotechnická

Czech Technical University in Prague
Faculty of Electrical Engineering

Dr. Tech. Ing. Radim Šára

Počítačové stereovidění

Computational Stereopsis

Summary

Computational stereopsis is a multidisciplinary field that tries to model the process of binocular image formation and formulate the algorithmic task of stereoscopic visual perception. The field has about 40-year history that started by studying the psychophysics of human stereovision. An artificial system that is able to perceive depth from passive vision has a great application potential. Today the methods are advanced enough to make computing three-dimensional geometric models of large real objects and scenes possible.

The algorithmic core of stereopsis is the matching problem. In this talk I will expose the problems stereopsis faces and I will try to formalize the matching problem in a such a way that perceptual illusions in complex scenes are avoided. This is where the traditional stereo methods fail. The proposed solution is based on stability rather than optimality and boils down to a graph-theoretic concept known as a directed graph kernel. I will generalize this concept to what I call strong sub-kernel, which is a structure that possesses more suitable properties for solving our problem.

Souhrn

Počítačové stereovidění je multidisciplinární obor, který se snaží modelovat proces binokulárního formování obrazu a formulovat algoritmickou úlohu stereoskopického vidění. Tento obor má zhruba čyřicetiletou historii a ve svých počátcích se zabýval studiem psychofyziky lidského stereovidění. Umělý systém, který je schopný vnímat hloubku z pasivního vidění, má velký aplikační potenciál. Dnes metody pokročily natolik, že je možné z obrazů automaticky vypočítat trojdimenzionální geometrické modely reálných objektů a scén.

Algoritmickým jádrem stereovidění je problém párování. V této přednášce ukáži problémy, kterým stereovidění čelí a pokusím se formalizovat problém párování tak, aby nevznikaly vizuální iluze ve složitých scénách. Právě v tomto případě tradiční metody selhávají. Navržené řešení je založeno na stabilitě, ne na optimalitě a vede k pojmu jádra orientovaného grafu, který je znám z teorie grafů. Zobecním tento pojem na tzv. silné sub-jádro, což je struktura, která má vlastnosti vhodnější pro vyřešení našeho problému.

Klíčová slova:	počítačové vidění, stereovidění, párování, stabilita
Keywords:	computer vision, stereopsis, matching, stability

České vysoké učení technické v Praze

Název:	Počítačové stereovidění
Autor:	Dr. Tech. Ing. Radim Šára
Počet stran:	23
Náklad:	150 výtisků

ISBN 978–80–01–04045–4

Contents

1 Introduction to Computational Stereopsis

Computer Vision is a field whose goal is interpreting images of real scenes. A defining characteristics of Computer Vision, as opposed to Image Processing, is that the scenes of interest are complex and hard to constrain and that one is not necessarily interested in interpreting the whole image. Another important characteristic is that interpretations must be made under uncertainty (multiplicity of likely interpretations) and noise (random noise due to sensor electronics and quantum nature of light but also non-random noise due to artifacts of low-level vision algorithms).

Because of the expected scene complexity, it is rarely possible to construct prior models that would be valid everywhere. But priors are needed to cope with uncertainty and noise. As a result, a standard low-level computer vision task uses only weak prior models. These tasks often involve matching (correspondence recognition) in general scenes. The ability to match features over a set of images constitutes a necessary basis for solving very many higher-level image interpretation problems.

Given two or more images, the goal of matching is to recognize which features in the target image(s) correspond to the features in the reference image. Since the general problem is difficult to tackle, two basic variants evolved over the course of history: The *Wide-Baseline Stereo Problem* and the *Semi-Dense Stereo Problem*. Their character and their solution covers most of the other cases as well. The Wide-Baseline Stereo Problem has applications in camera autocalibration, image stitching, recognition and image retrieval, visual tasks for robotic manipulation and navigation, range image registration, etc. Moreover, it is usually a necessary prerequisite to a successful solution of the Semi-Dense Stereo Problem. The Semi-Dense Stereo Problem has most applications in 3D modeling from images, in view synthesis, or in camera-based robotic obstacle avoidance.

1.1 Wide-Baseline Stereo Matching

The goal of Wide-Baseline Stereo in its simplest form is to recognize correspondences (matches) between a (relatively small) sets of points in the reference and the target images, as illustrated in Fig. 1. The images can be taken from very different viewpoints and possibly over long time periods. The usual first step involves finding a set of interest points in each image independently. These points are chosen to be well localized and stable under allowed image transformations [MTS+05]. A local image descriptor is then used to capture the content of the image neighborhood of each interest point [MS05]. The descriptor has to be invariant or at least insensitive to image deformations due to re-projection. Locality of descriptors is important for correspondence recognition in the presence of partial occlusion, time-induced image degradation factors, illumination changes, etc.

Let A and B be the interest point sets including their description in the reference and target image, respectively. The elements of $A \times B$ are the *putative correspondences* (pairs) p_i of wide-baseline stereo. The computational problem is to find the largest *partial* mapping $M\colon A \to B$ that has high probability and such that it satisfies additional constraints. The cardinality of the mapping is not known a priori.

A non-parametric condition on acceptable solutions M requires that each member of A and each member of B be matched at most once. This is called the *uniqueness constraint*. Assuming perspective camera, a parametric condition on acceptable solutions M has the following compact form:

$$\mathbf{y}(p_i)^\top \mathbf{F}\, \mathbf{x}(p_i) = 0 \text{ for all } p_i \in M, \tag{1}$$

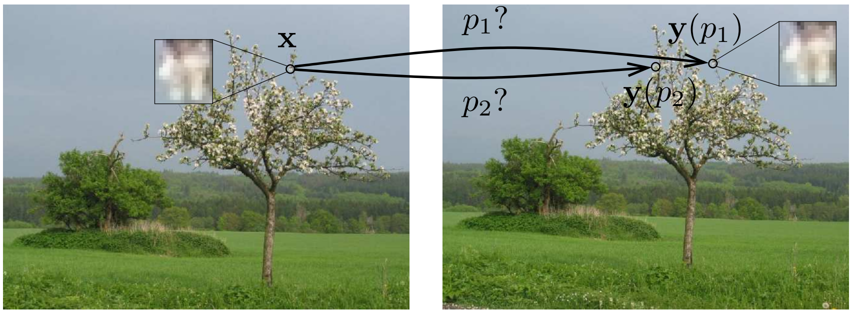


Figure 1: The goal of stereoscopic matching is to select a subset of putative (promising) correspondences p_i such that they are all consistent with a geometric model of a pair of cameras in a parametric form (1) (the parameters are unknown) and match similar local image descriptors (eg. image patches shown here).

where $\mathbf{x}(p_i)$, $\mathbf{y}(p_i)$ are image locations of an interest point in image A (B, respectively) expressed in homogeneous representation (it is a 3-vector) and $\mathbf{F}$ is a homogeneous 3×3 *fundamental matrix* of rank 2. The *epipolar constraint* (1) predicts that the corresponding point $\mathbf{y}(p_i)$ must lie on the line $\mathbf{F}\,\mathbf{x}(p_i)$ in the target image. The constraint has 7 independent parameters. The textbook [HZ03] gives a detailed description of geometric constraints related to projective cameras.

One possibility is to see this as a robust regression problem with additional constraint (uniqueness). Robustness is a mechanism that allows us rejecting putative correspondences that do not correspond to true correspondences. The standard, almost exclusively used solution to the wide baseline stereo problem is robust fitting of (1) by RANSAC [FB81, Tor00]. RANSAC (Random Sample Consensus) is a randomized method that samples 7-tuples of putative correspondences and retains those that show greater support in the set of putative correspondences (the consensus). The support is usually obtained by residual analysis of (1) (after normalization). RANSAC finds the solution with a given probability P. Insisting on $P = 1$ requires testing the support of all possible 7-tuples.

Such solution does not use all information available, namely the similarity of image neighborhoods of our interest points. Putative correspondences p_i that have similar descriptors (image neighborhoods) are more likely to form a solution. This observation is a basis for PROSAC, a variant of RANSAC that samples putative correspondences exhibiting image similarity more often than putative correspondences that do not [CM05]. RANSAC augmented with various speedups become quite wide-spread over its 25-year history [Ran06].

1.2 Semi-Dense Stereo Matching

In Semi-Dense Stereo, the interest points are the set of all image points. The goal is similar as above, with some simplifications that allow introducing additional models. Based on the fundamental matrix $\mathbf{F}$ obtained from the wide-baseline stereo correspondences it is possible to *rectify* (transform) the image domain so that the corresponding point in the target image is located on the same row as in the reference image [AH88, Har99, LZ99, GN01, MŠH04]. After the transformation, the parametric constraint (1) is no longer required. The image transformation not only means it is not necessary to search the whole image for a correspondence but also eases the use of some useful constraints: It has been observed [YP84] that in a wide class of scenes the left-to-right order in which interest

Figure 2: A binocular view of a scene of deep range (about 200 m) in which ordering is preserved (except for a little twig in the foreground tree).

points occur in the reference image is preserved in their respective matching points in the target image. This is called the *ordering constraint* which requires that the mapping M has a monotonicity property. Ordering constraint is violated by close and thin objects but in reality it holds in quite a broad class of scenes, as illustrated in Fig. 2.

Over the 40-year history of computational stereopsis, many different algorithms have been proposed that attempt to solve the semi-dense stereo problem, see [BBH03, SS02] for partial recent reviews (see also [DA89, Kos93] and Fig. 3). Among the least informed is the Winner-Take-All (WTA) algorithm. For each pixel in the reference image it computes image similarity with all possible corresponding pixels in the target image and accepts correspondence p maximizing this similarity. This does not work well, cf. Fig. 4, where the result is shown in the form of a *disparity map* in which color encodes relative depth. One of the failure reasons is the lack of a prior model. Such model is typically based on piecewise continuity of the solution. This is achieved by introducing the ordering constraint and a local *smoothness constraint*.

1.3 Occlusion

To be able to cope with even moderately complex scenes, semi-dense stereoscopic matching problem must include an *occlusion model*. This makes the matching problem significantly more difficult, as opposed to the case when no occlusions are allowed. Moreover, as we shall see shortly, there is no hope to make the model capture the whole world.

For simplicity, we assume only a single pair of cameras. We say a world point w (a point on a surface or in midair) is *ruled out* by a binocularly visible point p if either w is occluded by p in one of the cameras or if w is in front of p in one of the cameras. The situation is illustrated in Fig. 5. There are two fundamental types of (binocular) occlusion:

1. *Half-occlusion*: the set of surface points visible to both cameras rules-out all other world points. This case is illustrated in Fig. 5b.

2. *Mutual occlusion*: there are world points (in midair) that are not ruled-out by surface points visible to both cameras. This case is illustrated in Fig. 5c. Once the slit becomes wide enough for the background surface to enter the gray-shaded zone, the zone shrinks or disappears (especially if ordering is to hold, see later).

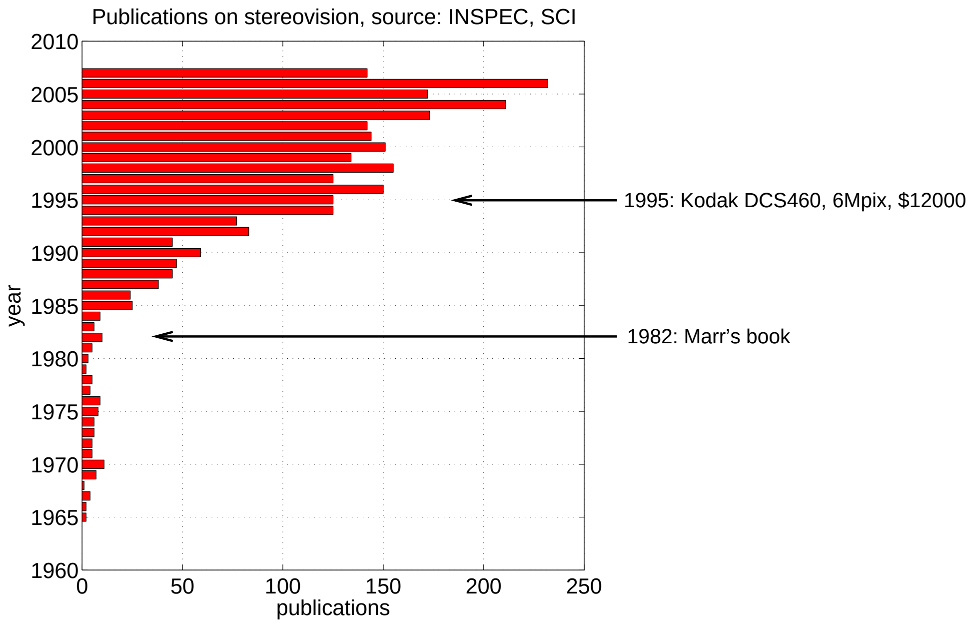


Figure 3: The annual number of publications on computational stereopsis. Sources: INSPEC, SCI Expanded (February 2008). Total: 2734 hits. Some of the bi-annual variations are attributed to the period of major Computer Vision conferences. Milestones that influenced interest in computational stereovision are Marr's book [Mar82] and the availability of digital cameras after 1995 that became widely affordable after about 2000. Results for 2007 are biased, not all publications of that year were included in the sources.

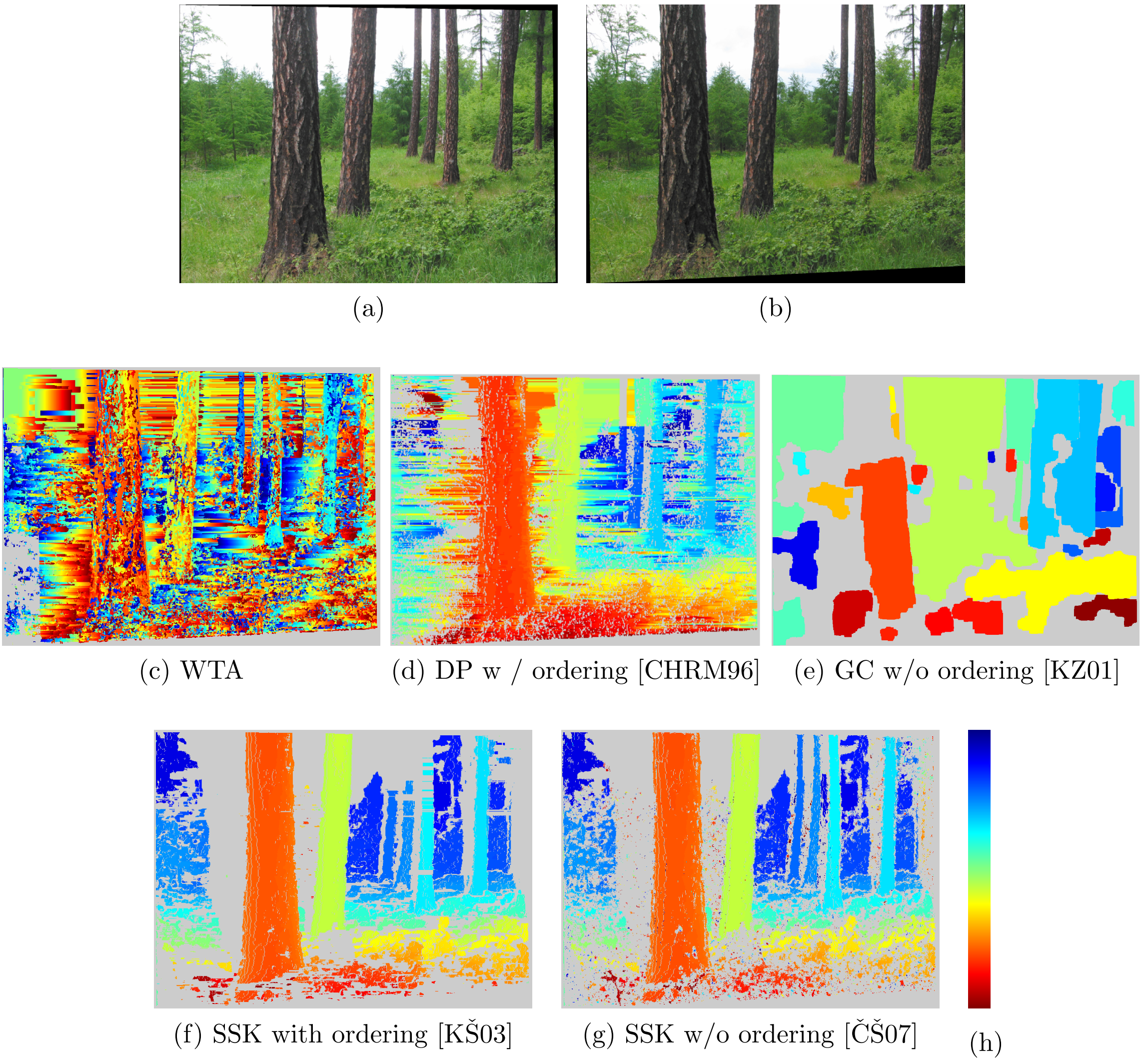
The scene in Fig. 4 has both kinds of occlusion. The region on the left of the foreground tree is half-occluded by the tree. The region between the foreground tree and the next tree is mutually occluded. In the latter case, all algorithms that do not model occlusion well, fail completely, cf. Figs. 4c-4e.

The uniqueness and ordering constraints are a half-occlusion models. Both are easy to incorporate into the matching task. The mutual occlusion is hard to formalize, which is perhaps the reason explicit attempts to formalize it have not been published in the literature so far.

Occlusion, mutual occlusion in particular, means we do not know a priori how large portion of the images can be interpreted as occluded or matched. It is clear that the unknown or unconstrained cardinality of the solution poses a serious problem in these tasks: The goal is not only to find a matching but also to determine which of the interest points are to be discarded. Of course one should discard as little as possible but prior knowledge useful for such disposal is hard if not impossible to obtain. With the exception of [GY05, Šár02, KZ02], none of the known algorithms models occlusion in a way that allows rejecting part of input data that is required here.

1.4 Repeated Appearance

Repeated or constant appearance is another difficult problem in stereopsis, especially when combined with occlusion: if the scene is a collection of small particles of equal color floating in the air, no local decision can determine which of the dots in the target image matches a particular dot in the reference image. The decision problem is somewhat easier if more than three cameras view the scene, especially if it is known that the scene consists of a surface visible to all cameras (i.e. when there are no occlusions) [BSK01, BSK03].

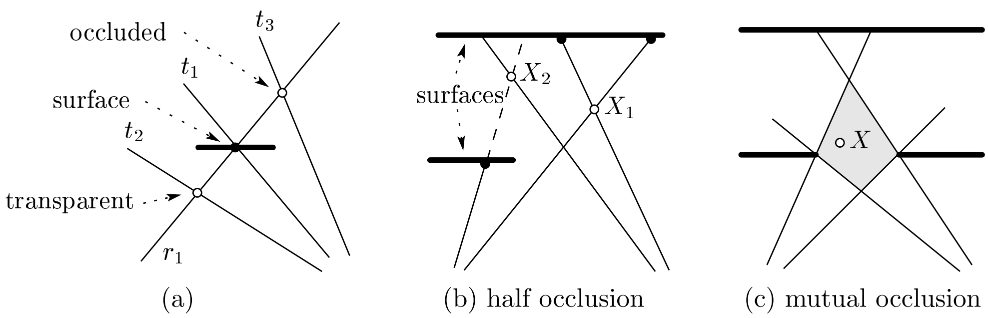


(a) (b)

(c) WTA (d) DP w / ordering [CHRM96] (e) GC w/o ordering [KZ01]

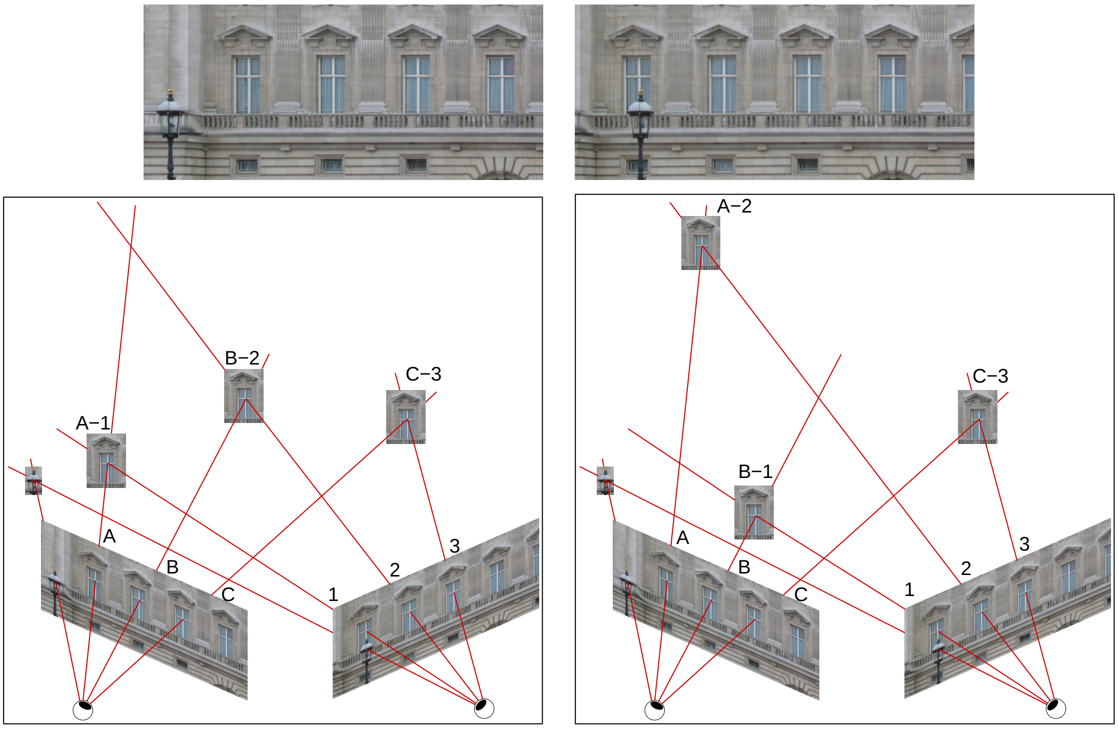
(f) SSK with ordering [KŠ03] (g) SSK w/o ordering [ČŠ07] (h)

Figure 4: Results of several semi-dense stereo algorithms on a non-trivial image pair. Color (h) codes left-image disparity: large disparity of close objects is red, small disparity of distant objects is blue, regions where no matches were assigned are gray.



(a) (b) half occlusion (c) mutual occlusion

Figure 5: Occlusion: The surface point at the intersection of rays r_1 and t_1 (solid) occludes a world point at the intersection (r_1, t_3) and implies the world point (r_1, t_2) is transparent, hence (r_1, t_3) and (r_1, t_2) are ruled-out by (r_1, t_1) (a). In half-occlusion, every world point such as X_1 or X_2 is ruled out by a binocularly visible surface point (b, solid dots). In mutual occlusion this is no longer the case (c, gray region).



matching 1 matching 2

Figure 6: Structural ambiguity in stereopsis in the case of repeated appearance. Unless we know all windows share a common façade plane, we cannot decide which matching is correct. But planarity assumption is an unrealistic model even in this scene (consider the candelabrum).

Also, if a portion of the scene has a constant appearance (consider a perfectly white wall), unique solution does not exist regardless of the number of cameras viewing it.

Figure 6 helps illustrate the repeated appearance problem. A way out is *scene understanding:* if we can segment-out the image region that corresponds to the façade (based on the knowledge of the world of façades and their images) and employ the knowledge façades are (almost) flat, we can match the images unambiguously. In fact, in this particular case, some simpler approach would work as well: ordering constraint disambiguates the solution for the façade (but is violated by the candelabrum, a minor element in the scene). Alternatively, the minimum description length (MDL) principle [Ris78] could disambiguate the solution as well. But there is always a sufficiently complex scene where MDL fails. Hence, a low-level problem (matching) becomes a scene interpretation problem in its full complexity. This is a typical example illustrating the complexity of vision. In this work we will avoid dependencies between low-level and high-level vision by introducing *robustness*. See Sec. 2. In our ongoing collaborative effort [eTR09] we try to couple high-level scene interpretation with low-level image interpretation tasks.

1.5 Directional Reflectance

Another source of ambiguity are image descriptors themselves. This is because surface reflectance is directionally dependent, which makes the descriptors not directly comparable [NIK91]. The simplest illustration of the problem are specularities, see Fig. 7: they tend to 'shift' on the surface with the change of the viewpoint. The motion strongly coupled with differential surface properties [Kv80].

Even non-specular surfaces have directional reflectance (see [NIK91] for a review).

Figure 7: Directional surface reflectance makes local image descriptors incomparable. Here, the illuminant did not change as the camera moved.

Moreover, reflectance is a material-dependent property. A common brute-force approach to the problem is normalization of the image descriptor so that it is invariant to some class of allowed image transformation (eg. by linear normalization or Census Transform [ZW94]) or using an invariant correlation statistics (eg. normalized cross-correlation coefficient invariant to linear transformations or rank correlation invariant to all monotonic transformations [BN96]). The disadvantage of such approach is that the descriptor is losing its discriminability potential (the ability to discriminate among many different image locations).

The only known way to circumvent the directional reflectance problem is *Helmholtz stereopsis* [ZBK02]. It employs reciprocity of the reflectance law [vH89] following from time-reversal invariance [SWL98]. A Helmholtzian stereoscopic setup is possible when lights and cameras are co-located. A stereoscopic image pair is obtained by (1) switching on light L_1 in Location 1 and recording image I_2 in camera C_2 in Location 2 and then (2) by switching on L_2 and recording image I_1 in C_1. The pair I_1, I_2 has the remarkable property that image values of corresponding pixels are the same, up to a scalar constant that depends on depth and local surface normal orientation [MKZB01].

1.6 Two Alternative Problem Formulations

The arguments reviewed so far suggest the complexity of the stereoscopic matching problem is probably beyond reasonable formalization if the task is considered in full complexity. To approach the problem, simplifications must be made.

One way to avoid the multiplicity of interpretations is to use regularization. This requires a prior model and will be discussed shortly in Sec. 1.6.1. Another way is to employ *stability*. This is shortly discussed in Sec. 1.6.2.

1.6.1 Regularized Matching Problem

The regularized semi-dense stereo matching problem under smoothness (and optionally ordering) constraint may be posed as a Bayesian decision task with unit cost that leads to maximum aposteriori probability algorithm (MAP). Formalization of this task requires some care. First, to properly model continuity, we need the concept of *disparity space*. It is nothing but the discrete set of putative correspondences $P(r) = A(r) \times B(r)$, one per rectified image row r and stacked to form a 3-dimensional array P. Putative hypotheses become triples (r, i, j), where r is the common image row, i is a column index in the left image and j is a column index in the right image. We say two correspondences p, q in disparity space P are neighbors if they are neighbors in either image, i.e. if they fall on neighboring optical rays in one or the other camera.

Assuming suitable probability distributions the MAP task becomes one of minimizing discrete energy

$$M^* = \arg\min_{M \in \mathcal{M}} \; \alpha_0 \sum_{p,q \in N(M)} V_{12}(p,q) + \sum_{p \in M} V_1(p), \tag{2}$$

where $\alpha_0 > 0$ is the relative weight of the prior model, $\mathcal{M}$ is the set of all one-to-one matchings, possibly conforming to additional constraints like ordering, V_1 is energy related to image similarity for putative correspondence p, and V_{12} is related to prior continuity or smoothness model in which $N(M)$ is the set of all pairwise disparity space neighbors included in matching M. To avoid degeneracy when solving the problem (2), the problem must be formulated so that all matchings from $\mathcal{M}$ have the same cardinality. This is discussed in the main thesis [Šár07].

Unfortunately, the solution to the general problem when the prior energy V_{12} includes neighbors across the image row coordinate r is not known. In the known attempts to formulate the problem the images play asymmetric role, eg. [KZ01], the elements of the set of putative solutions $\mathcal{M}$ do not have all the same cardinality which leads to degeneracies and artifacts, eg. [CHRM96, Gim99], the formalization is not able to incorporate uniqueness constraint, eg. [KZ01, BVZ01], or ordering constraint, eg. [KZ01, BVZ01] (let alone model mutual occlusion), and/or discontinuities are not allowed at all, eg. [SFS04].

We may restrict ourselves to the case when there are no neighbors with respect to the first coordinate in disparity space (across image rows) and ordering is strictly enforced. In such case the problem reduces to finding a minimum-cost path in an acyclic directed graph, where both node and arc costs are non-negative. This is a well-known problem. Details are given in the main thesis [Šár07]. Results are shown in Fig. 4d.

The horizontal streaks in the disparity map in Fig. 4d are not caused by the lack of regularization in the vertical direction. They are caused by the fact that there are multiple optima of (2), especially if there are texture-less areas and/or if the prior model (ordering) is violated. In such case a small noise causes the solution to be *unstable.* The instability can be formalized, which is the main topic of the thesis [Šár07].

Attempts to incorporate isotropic continuity/smoothness prior have been made but the problem was found of non-polynomial complexity. Generally, the task belongs to *consistent labeling problems* [FS00, KZ04]. Fig. 4e shows a result using the approximation algorithm described in [KZ01]. I believe the patchiness is mostly due to an insufficient approximation power of the algorithm and the fact it uses smoothness constraint but neither uniqueness nor ordering constraints. In this case the instability is manifested by entire patches rather than streaks. Gong and Yee-Hong [GY05] tried to address the instability within the framework of energy minimization.

1.6.2 Stable Matching Problem

We have seen in Fig. 4 that simple algorithms like WTA do not work because of lack of prior models but that the MAP algorithm making use of such models suffer from artifacts, most often from false positive 'illusions.' Such illusions are then propagated up the image interpretation process where it is difficult to suppress them. An example application field where this is a serious problem is 3D scene reconstruction.

Results in Figs. 4f and 4g show disparity maps computed by algorithms that directly address the instability mentioned at the end of previous subsection. The algorithms are described in great detail in the main thesis [Šár07]. The theory will be sketched in Sec. 2.

1.9	1.3	1.2	2.0
0.2	0.3	0.9	1.3
1.1	0.3	0.1	1.5
1.0	1.2	0.2	1.1

(a) M_1

1.9	1.3	1.2	2.0
0.2	0.3	0.9	1.3
1.1	0.3	0.1	1.5
1.0	1.2	0.2	1.1

(b) M_2

Figure 8: Min-cost solution M_1 to a perfect bipartite matching problem of cost 2.7 that is not absorbing (a) because of the red unmatched element. An absorbing solution M_2 to the same problem with the same total cost of 2.7 and absorbance margin of 0.1 (b).

The point to be made here is that the energy minimization algorithm is unstable because it does not guarantee *absorbance*. I will demonstrate it on a simple example. We first need some notation: Given an element $p = (i, j)$ in a matching table like the one in Fig. 8a we denote the set $X(p)$ as

$$X(p) = \Big\{(i, k) \cup (l, j), k \neq j, l \neq i\Big\}, \quad p = (i, j), \tag{3}$$

which is also illustrated in Fig. 9a. Note that a subset M of the matching table is a matching iff for each $p \in M$ it holds $X(p) = \emptyset$. Note that if M is a perfect matching (i.e. one that assigns a single match to every column and every row of the table), then each $q \notin M$ has two elements in $X(q) \cap M$.

Suppose now we have a matching table with non-negative real entries, like the one in Fig. 8 and our goal is to find a minimum-cost perfect bipartite matching. The result in Fig. 8a shows a solution M_1 to the problem in gray table cells. The cost of the solution is 2.7. Another solution M_2 of the same cost is shown in Fig. 8b. The latter solution M_2 is stable since it is absorbing in the following sense: every *unmatched* element $q \notin M_2$ of the table (in Fig. 8b) has at least one element in $X(q) \cap M_2$ of a *strictly lower cost* than q. In fact, in this example the *absorbance margin* is 0.1: even if any combination of the unmatched elements decreased their costs by 0.1 (or increase their cost by any amount), the solution M_2 would still be both optimal and absorbing.

The solution in Fig. 8a is not absorbing because it does not absorb the element of cost 0.1 at position 3,3 of the table (shown in red). A very small change of its cost (to $0.1 - \epsilon$ for some small positive ϵ) causes the optimal solution M_1 switch to M_2. The stable solution does not necessarily have the same cost as the optimal solution as the example might suggest. Generally, its cost is just sub-optimal and the level of sub-optimality can be characterized, see Sec. 2.2.

The problem with stability of the solution is not unique to just bipartite matching problems. It pertains to all min-sum and max-sum problems.

There is an algorithm for finding (strongly) absorbing (stable) solutions. This is one of the main contributions presented in the thesis [Šár07]. The algorithm is sketched in Algs. 1 and 2. Interestingly, stability is related to robustness: If we increase the stability margin, then, upon proper generalization of stability, eg. by Def. 1, the solution is no longer complete and part of images remain uninterpreted (rejected). The rejection occurs in case of local violation of prior models (like ordering), which is exactly the behavior we need in a world which cannot be explained by a single universal model because the complexity of such model would be overwhelming, as discussed at the end of Sec. 1.4.

Standard matching methods that are based on classical discrete energy minimization cannot cope with the problem without explicitly introducing a special label 'rejected,'

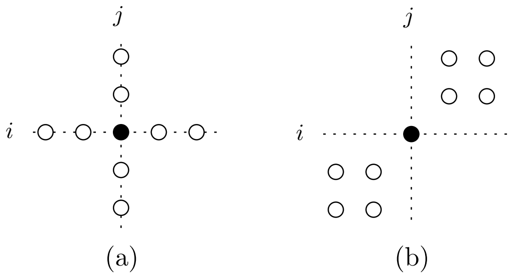


Figure 9: Matching table representation of the sets $X(p)$ (left) and $F(p)$ (right). The solid vertex is the pair $p = (i, j)$.

which necessarily destroys any structural properties of the label set [FS00, KZ04]. As a result, the matching (correspondence recognition) problem becomes NP.

My program over the last several years has been exploring a formal definition of stability to achieve robustness in the sense of selective data interpretation. The goal has been to achieve algorithmic simplicity as well. The thesis [Šár07] shows that the intuitive definition of stability introduced in the preceding subsection is related to *digraph kernels*. The concept must be slightly modified to be more useful in computer vision tasks. The thesis shows that in fact a variety of computer vision problems can be solved based on the stability principle. I sketch the formalization in the rest of this section and then overview some theoretical results in Sec. 2.

As before, A, B are two sets of *participants* of the matching game. The set $V \subseteq A \times B$ are putative correspondences. One can imagine A, B to be the sets of optical rays (casted by the aforementioned interest points) in the reference and target cameras, respectively, and $A \times B$ to be the set of all their mutual spatial intersections, as in Fig. 6. Our goal is to find the best partitioning of V to three subsets: matched M, uninterpreted U and ruled-out R (occluded or transparent). We will construct a simple graph $\mathcal{G} = (V, E)$ over the set V as follows. If there are two vertices $v_1, v_2 \in V$ that cannot be members of the solution simultaneously, we add edge (v_1, v_2) to E. For instance, since the matching is to be one-to-one (due to occlusion or transparency), each participant can be matched at most once. Hence, the set of neighbors in $\mathcal{G}$ of a vertex p includes the set $X(p)$ introduced in (3) and Fig. 9a.

Other constraints can be included as well. If ordering is assumed, the resulting matching M must be monotonic and the set of all neighbors for $p = (i, j)$ includes the set $F(p)$ of all pairs (k, l) such that $k < i$ and $l > j$ or $k > i$ and $l < j$. In the matching table representation the element (i, j) is connected to all elements in two opposite quadrants, see Fig. 9b.

Problems involving parametric constraints can be formalized as well. An example is the WBS problem. Let the set of parametric constraints have m parameters and let a single pair (i, j), $i \in A$, $j \in B$ remove d degrees of freedom from the constraint set. E.g., the constraint (1) has $m = 7$ parameters and each point correspondence removes $d = 1$ degree of freedom. In this case we proceed as follows: the participants of the matching game are the sets A^4, B^4, i.e. all interest point quadruples. A pair of octuples $\{i_{11}, i_{12}, i_{13}, i_{14}; j_{11}, j_{12}, j_{13}, j_{14}\}$, $\{i_{21}, i_{22}, i_{23}, i_{24}; j_{21}, j_{22}, j_{23}, j_{24}\} \in A^4 \times B^4$ are connected by edge in E if the set of points $\mathbf{x}_{i_{1k}}$ and $\mathbf{x}_{j_{2k}}$ do not satisfy (1) for all $k = 1, 2, 3, 4$. We need strictly more than $\frac{m}{d}$ correspondences which requires participant sets to be r-tuples A^r, B^r, where r is the smallest integer strictly greater than $\frac{m}{2d}$. The growth of

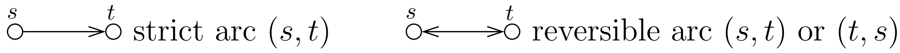


Figure 10: Strict and reversible arc of an oriented graph $(\mathcal{G}, \omega)$.

the dimension r of the problem can be avoided by more rich local image features: For instance, if ellipses are used then $d = 2$ [HZ03], and the participant sets are just pairs from A^2, B^2.

Edges due to uniqueness or ordering constraints are as easy to add to the graph over the vertex set $A^r \times B^r$ as above. To summarize, the graph is $\mathcal{G} = (V, E)$ where $V = A^r \times B^r$ captures the structure of all geometric and parametric constraints of the given problem. It is important to observe that independent vertex sets of graph $\mathcal{G}$ represent the set of feasible solutions. This is the set on which we will be selecting the best solution, given data and prior knowledge.

Let $V(\mathcal{G})$ denote the vertex set of graph $\mathcal{G}$. Let $e(v)$ be a closed real interval for every $v \in V(\mathcal{G})$. We call it the *evidence interval* here. The interval captures the probability $p(v \in M \mid \mathbf{z})$, i.e. the probability that v is a correct match given measurement $\mathbf{z}$. The width of the interval represents our uncertainty on the true value of $p(v \in M \mid \mathbf{z})$ due to data noise, known bias, approximation, and/or other reasons. The width of the interval can be adjusted by a user-selected confidence parameter.

If the intervals are $[0, 1]$ for all $v \in V(\mathcal{G})$, data is totally uninformative and we are expecting an empty solution. The narrower the intervals the greater fraction of data is expected to be interpreted (unambiguously).

We say a vertex $t \in V(\mathcal{G})$ is a *competitor* to vertex $s \in V(\mathcal{G})$ if s and t are connected by an edge in $\mathcal{G}$ and $\max e(t) > \min e(s)$ (the $e(t)$ is greater than $e(s)$ or the intervals overlap). We say an independent vertex set M of $V(\mathcal{G})$ is *stable* if every vertex $q \notin M$ has at least one of its competitors in M, in other words, if there is a reason for such q to be ruled-out. We can obtain a purely graph-theoretic representation of the matching problem as follows: the underlying graph $\mathcal{G}$ is as before. We construct orientation ω of the edges of $\mathcal{G}$ as follows: if $\{s, t\} \in E(\mathcal{G})$ and $\max e(t) > \min e(s)$ we orient the arc from s to t. If $\{s, t\} \in E(\mathcal{G})$ and the intervals $e(t)$ and $e(s)$ overlap we orient the arc bidirectionally, see Fig. 10. We call the result an *interval orientation* of the underlying graph to distinguish it from a general orientation of the graph. Interval orientations have a number of important properties.

To summarize, the pair $(\mathcal{G}, \omega)$ is a digraph in which some arcs can have both orientations. The stable set M of $(\mathcal{G}, \omega)$ is then an independent vertex subset such that each vertex $q \notin M$ has a successor in M. This structure is known as a directed graph *kernel* [vNM44, BG03].

The stable sets (kernels of $(\mathcal{G}, \omega)$) are our prospective solutions. They are not yet 'stable enough,' consider the example in Fig. 11b, where it is not possible to choose which solution (green or red) is better. We feel the red solution appears more stable because it is not influenced by choosing a definite orientation for the top bidirectional arc. We will now formalize strong stability that captures the difference.

We say an arc $(s, t) \in \omega$ is *strict* if $(t, s) \notin \omega$. Otherwise it is called *reversible* (rather than bidirectional). See Fig. 10. We say t is a successor of s if there is arc (s, t) and we say t is a strict successor of s if there is arc (s, t) but not (t, s) in $(\mathcal{G}, \omega)$. To introduce robustness to stable sets, we define strong sub-kernel as follows:

Definition 1 (SSK)**.** *Let $(\mathcal{G}, \omega)$ be an oriented graph. An independent vertex subset $S \subseteq V(\mathcal{G})$ is a strong sub-kernel if every successor of each $v \in S$ has a strict successor*

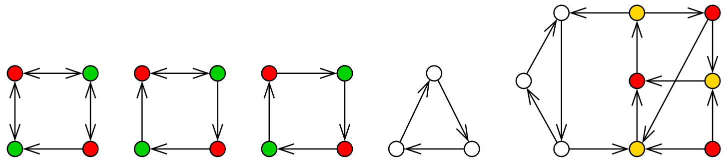


Figure 11: Several orientations with their kernels and maximum strong sub-kernels (SSK). Kernels in (a)–(d) are distinguished by color. Number a/b indicates the orientation has a kernels and b maximal SSKs. The orientation in (e) has no kernel but has a single maximal SSK (red). Only (a) is an interval orientation.

in S.

Fig. 11 shows several examples of maximal strong sub-kernels (SSK) in several oriented graphs: (a) and (d) have no SSK, (b) has one SSK (red), (c) has two SSKs (red, green), and (e) has an SSK (red) despite the fact it has no kernel.

Let us check if SSK has the desired behavior. If data is not informative, all arcs are reversible and the solution is empty. This was intended. If data is in contradiction with the model (represented by the underlying graph $\mathcal{G}$) then even in the absence of evidential uncertainty, part (or all) of the graph gets rejected, as the example in Fig. 11e shows. In the case of Fig. 11e we have partitioned the vertex set to three subsets: matched (red), ruled out (gold) and uninterpreted (white).

The prefix sub- in 'strong sub-kernel' has been chosen to indicate incompleteness: the SSK is no longer a maximal independent set. Using standard terminology, a *maximal* SSK is not extendible to a larger SSK and a *maximum* SSK then has the largest cardinality of all maximal SSKs. Note than incompleteness is necessary to obtain robustness. Maximality of SSK implies minimality of the uninterpreted vertex subset.

As shown in the thesis [Šár07] the problem of finding an SSK in interval orientation is solvable by a very simple algorithm:[1]

Algorithm 1 (Sink Stripping)**.**

Input: *An interval-oriented graph* $(\mathcal{G}, \omega)$*.*
Output: *Maximum strong sub-kernel* S*.*
Procedure:

1. *Initialize* $S := \emptyset$*.*
2. *If there is no sink in* $\mathcal{G}$*, terminate and return* S*.*
3. *Find a sink* $s \in V(\mathcal{G})$*.*
4. *Add* s *to* S*.*
5. *Remove* s *and all its predecessors* $P(s)$ *from* $\mathcal{G}$*.*
6. *Go to Step 2.*

This basic version of the Sink Stripping algorithm has worst-case complexity of $O(\alpha\, n)$, where α is the independence number of $\mathcal{G}$ and n is the number of its vertices. For interval

[1]Sink is a vertex with no successor. Isolated vertex is a sink.

orientations of graphs that occur in matching under uniqueness (and, optionally, ordering) there is a faster algorithm of worst-case complexity $O(n \log n)$ [Šár07].

Finding sinks in an oriented graph may be difficult, especially if the graph is dense and not explicit. A better algorithm for the class of orientations to which interval orientations belong is described the thesis [Šár07].

The class of oriented graphs for which Alg. 1 works is broader than interval orientations. Unfortunately, there is no known algorithm for a general orientation of the underlying graph. A solvable class will be discussed in the next section.

2 Strong Sub-Kernels in Oriented Graphs

In this section we briefly review the main theoretical results of the thesis [Šár07]. We consider Definition 1 in which $(\mathcal{G}, \omega)$ will be a general orientation of a general underlying graph $\mathcal{G}$. We have seen several examples of SSKs of general oriented graphs in Fig. 11.

The thesis [Šár07] considered the following questions:

1. What is the class of orientations of general underlying graphs that have at most one maximal SSK?

2. Are there any notable properties of the class found, especially those related to robustness and optimality?

3. What are the algorithms?

I will briefly review the main results in the subsequent three subsections. The reader is referred to [Šár07, Šár06, BŠ07] for proofs and detailed discussion.

2.1 Graphs with at Most One SSK

We say a subgraph $\mathcal{C}$ is an *EC-subgraph* of an oriented graph $(\mathcal{G}, \omega)$ if it is induced in $(\mathcal{G}, \omega)$ by an even circuit (circuit of even length). The class is then characterized by the following theorem:

Theorem 1 (No. 23 in [Šár07])**.** *Let every EC-subgraph of oriented graph $(\mathcal{G}, \omega)$ have at most one maximal SSK. Then $(\mathcal{G}, \omega)$ has at most one maximal SSK.*

There is a number of ways to achieve a single SSK per EC-subgraph. For instance, there is a structural condition in the underlying graph $\mathcal{G}$ ensuring every even circuit $\mathcal{C}$ in $\mathcal{G}$ has a an *even chord* (ie. one connecting two even (or two odd) vertices in $\mathcal{C}$). Alternatively, we may restrict the orientation ω of $\mathcal{G}$ so that every even circuit in $\mathcal{G}$ has a reversible arc (called sometimes a *pseudo-chord* [BKW98]). It is easy to verify circuits with even chords or pseudo-chords have at most one SSK.

Interval-oriented graphs fall in the class required in Theorem 1. More specifically,

Theorem 2 (No. 15 in [Šár07])**.** *Every circuit in interval-oriented graph has at least two consecutive reversible arcs.*

Interval orientations also have some stronger properties that will be hinted in the next section.

2.2 Robustness and Optimality

Strong sub-kernels are stable with respect to uncertainty increase or reduction in oriented graph $(\mathcal{G}, \omega)$ in the following sense:

Theorem 3 (Lemma 19 in [Šár07]). *Let $(\mathcal{G}, \omega)$ have no SSK. Then every $(\mathcal{G}, \omega')$ constructed from $(\mathcal{G}, \omega)$ by making an arbitrary subset A of arcs in ω reversible has no SSK.*

Theorem 4 (Lemma 20 and Corollary 14.1 in [Šár07]). *Let S be an SSK in $(\mathcal{G}, \omega)$, let $(\mathcal{G}, \omega')$ be constructed from $(\mathcal{G}, \omega)$ by making an arbitrary subset of reversible arcs irreversible and let S' be the maximal SSK in $(\mathcal{G}, \omega')$. Then S is an SSK in $(\mathcal{G}, \omega')$. Moreover, if ω is EO then $S \subseteq S'$.*

EO orientations are such that every even circuit has an even and an odd reversible arc. Interval orientations are EO [Šár07]. In interval orientations, Theorem 4 implies that an SSK for a given set of intervals $e(p)$ is the intersection of all SSKs, each obtained for some choice of intervals $e'(p)$ such that $e'(p) \subseteq e(p)$ for all vertices p in $(\mathcal{G}, \omega)$ [Šár06].

This behavior is related to robustness since widening the intervals is essentially a way of increasing stability margins discussed in Sec. 1.6.2. Wide $e(v)$ is a safeguard against error or bias in the estimate of $p(v \in S \mid \mathbf{z})$ or represents our inability to provide its accurate value based on data collected so far. Controlling the widths of the intervals provides a way to control the 'degree of stability.' Practical experience shows that SSK indeed exhibits robust behavior by not explaining data that contradicts prior model. An example is shown in Fig. 4f: Some of the trees in the background appear 'chopped off.' This is because they locally violate ordering that has been assumed. If we remove ordering constraint, we obtain the result in Fig. 4g. The price is, of course, more errors (numerous isolated odd-colored pixels), since the model became weaker.

Let us now turn our attention to optimality. Of course, robustness and optimality in the sense of energy optimization are contradictory requirements. The following two theorems characterize the remaining weak optimality of SSK.

Let $P(p)$ be the set of all predecessors of vertex p in $(\mathcal{G}, \omega)$. Let $Q(p)$ be an independent[2] subset of $P(p)$ and $\mathcal{Q}(p)$ be the set of all such subsets $Q(p)$. We denote $\underline{e}(p) = \min e(p)$ and $\overline{e}(p) = \max e(p)$.

Theorem 5 (Weak Optimality I [Šár06]). *Let $(\mathcal{G}, \omega)$ be an interval-oriented graph and let M be a SSK which is also a maximal independent vertex set of $(\mathcal{G}, \omega)$. If there is a sequence of sinks s chosen in Step 3 of Alg. 1 such that each s satisfies*

$$\underline{e}(s) \geq \max_{Q \in \mathcal{Q}(s)} \sum_{q \in Q} \overline{e}(q) \tag{4}$$

at the moment of its selection in Step 3, then M also maximizes the cost sum over all possible independent vertex subsets $\mathcal{M}$ of $(\mathcal{G}, \omega)$,

$$M = \arg\max_{K \in \mathcal{M}} \sum_{p \in K} \overline{e}(p). \tag{5}$$

The converse uses a much weaker condition: It requires *each vertex* of $V(\mathcal{G})$ to satisfy (4):

Theorem 6 (Weak Optimality II [Šár07]). *Let $(\mathcal{G}, \omega)$ be an interval-oriented graph, M be a solution to (5) and let each vertex $p \in V(\mathcal{G})$ satisfy (4) in $(\mathcal{G}, \omega)$. Then M is a maximum SSK in $(\mathcal{G}, \omega)$.*

[2] An independent vertex subset in subgraph induced by $P(p)$ in $(\mathcal{G}, \omega)$.

2.3 The Octopus Algorithm

Let every EC-subgraph of $(\mathcal{G}, \omega)$ have a reversible arc (we say it is ECRA). Then the following algorithm finds the unique maximal SSK in $(\mathcal{G}, \omega)$:

Algorithm 2 (The Octopus Algorithm)**.**

Input: *An ECRA oriented graph* $(\mathcal{G},\, \omega)$*, in which* A *are arcs in* $(\mathcal{G},\, \omega)$*.*
Output: *Maximum strong sub-kernel* S*.*
Working data-structures:
G *vertices with a strict successor which is not in* R
R *vertices with a successor which is not in* G

Procedure:

1. *Initialize* $R := \emptyset$*.*
2. *Reset* $G := \emptyset$*.*
3. *For every irreversible arc* $e^* = (v_1,\, v_2) \in A$*,* $(v_2,\, v_1) \notin A$ *do:* *if* $v_2 \notin R$ *then* $G := G \cup \{v_1\}$*.*
4. *For every arc* $e = (v_1,\, v_2) \in A$ *(reversible or not) do:* *if* $v_2 \notin G$ *then* $R := R \cup \{v_1\}$*.*
5. *Repeat Steps 2–4 until there is no change in R.*
6. *If* $G = \emptyset$ *or* $G \subseteq R$ *terminate and return* $S = V \setminus R$*.*
7. $R := R \cup G$*, go to Step 2.*

The algorithm is essentially a message propagating procedure which waits until the R-G coloring of graph vertices reaches an equilibrium. The messages are initially induced by all arcs in $(\mathcal{G}, \omega)$.

The worst-case time complexity of Alg. 2 is $O(m\,\alpha)$, where m is the number of edges in $\mathcal{G}$ and α is the independence number of $\mathcal{G}$. It is easy to see $|S| \leq \alpha$, where $|S|$ is the cardinality of the solution.

3 Conclusions

The principle of stability gives new robust algorithms for solving matching (and some other) problems in computer vision. Unlike the popular randomized robust algorithm RANSAC [FB81] and its variants [Ran06] the new algorithms are deterministic, of low-order polynomial complexity and are fast in practice, which means they can be used in large problems like dense stereo matching. Our experience suggests algorithms based on stability tend to find more accurate solutions than RANSAC. A certain disadvantage of the stable matching algorithms is that all the evidence intervals must be available prior to running the algorithm. Further research is needed to overcome the problem, at least partially.

So far, the stability principle has successfully been applied to several matching problems in computer vision: dense stereoscopic matching [Šár02, KČŠ03], including a very fast on-line version of the algorithm [ČŠ07], range image registration [ŠOS05, SSŠ07], and homography estimation [BŠ07].

There are a number of interesting open problems. The class of oriented graphs that have at most one maximal strong sub-kernel is not yet explored well enough: the necessary and sufficient condition for the existence of at most one SSK is unknown. It is not known if the Octopus algorithm finds a maximum kernel in other oriented graphs than the ECRA graphs, for instance, does it work for the class of graphs defined in Theorem 1? Is there a weighted version of the strong sub-kernel? Can we find stable solutions to energy minimization problems? There are many more open questions, some of which are stated in the thesis [Šár07].

References

[AH88] N. Ayache and C. Hansen. Rectification of images for binocular and trinocular stereovision. In *Proceedings of 9th International Conference on Pattern Recognition*, volume 1, pages 11–16, October 1988.

[BBH03] M. Z. Brown, D. Burschka, and G. D. Hager. Advances in computational stereo. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(8):993–1008, August 2003.

[BG03] E. Boros and V. Gurvich. Perfect graphs, kernels, and cores of cooperative games. Rutcor Research Report RRR 12–2003, Rutgers Center for Operations Research, Rutgers University, April 2003.

[BKW98] O. V. Borodin, A. V. Kostochka, and D. R. Woodall. On kernel-perfect orientations of line graphs. *Discrete Mathematics*, 191(1-3):45–49, September 1998.

[BN96] D. N. Bhat and S. K. Nayar. Ordinal measures for visual correspondence. Research Report CUCS-009-96, Department of Computer Science, Columbia University, New York, February 1996.

[BŠ07] M. Bujňák and R. Šára. A robust graph-based method for the general correspondence problem demonstrated on image stitching. In *Proceedings of the International Conference on Computer Vision*, page 8, 2007.

[BSK01] S. Baker, T. Sim, and T. Kanade. A characterization of inherent stereo ambiguities. In *Proceedings of the International Conference on Computer Vision*, volume 1, pages 428–435, July 2001.

[BSK03] S. Baker, T. Sim, and T. Kanade. When is the shape of a scene unique given its light-field: A fundamental theorem of 3D vision? *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(1):100–109, January 2003.

[BVZ01] Y. Boykov, O. Veksler, and R. Zabih. Fast approximate energy minimization via graph cuts. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(11):1222–1239, November 2001.

[CHRM96] I. J. Cox, S. L. Higorani, S. B. Rao, and B. M. Maggs. A maximum likelihood stereo algorithm. *Computer Vision and Image Understanding*, 63(3):542–567, May 1996.

[CM05] O. Chum and J. Matas. Matching with PROSAC – progressive sample consensus. In *IEEE Conference on Computer Vision and Pattern Recognition*, volume 1, pages 220–226, 2005.

[ČŠ07] J. Čech and R. Šára. Efficient sampling of disparity space for fast and accurate matching. In *Proceedings of CVPR Workshop Towards Benchmarking Automated Calibration, Orientation and Surface Reconstruction from Images*. Omnipress, 2007.

[DA89] U. Dhond and J. Aggarwal. Structure from stereo—a review. *IEEE Transactions on Systems, Man, and Cybernetics*, 19(6):1489–1510, 1989.

[eTR09] eTRIMS – E-training for interpreting images of man-made scenes, 2007-2009. STREP Project No. FP6-IST-027113. http://www.ipb.uni-bonn.de/projects/etrims/.

[FB81] M. A. Fischler and R. C. Bolles. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 24(6):381–395, June 1981.

[FS00] B. Flach and M. I. Schlesinger. A class of solvable consistent labelling problems. In *Advances in Pattern Recognition. Proceedings of Joint IAPR International Workshops SSPR and SPR 2000*, volume LNCS 1876, pages 462–471, 2000.

[Gim99] G. Gimel'farb. Stereo terrain reconstruction by dynamic programming. In Jähne and et al., editors, *Handbook of Computer Vision and Applications*, volume 2, chapter 18, pages 505–530. Academic Press, 1999.

[GN01] J. Gluckman and S. K. Nayar. Rectifying transformations that minimize resampling effects. In *IEEE Conference on Computer Vision and Pattern Recognition*, volume 1, pages 111–117, December 2001.

[GY05] M. Gong and Y.-H. Yang. Unambiguous stereo matching using reliability-based dynamic programming. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(6):998–1003, June 2005.

[Har99] R. I. Hartley. Theory and practice of projective rectification. *International Journal of Computer Vision*, 35(2):115–127, 1999.

[HZ03] R. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision.* Cambridge University Press, 2nd edition, 2003.

[KČŠ03] J. Kostková, J. Čech, and R. Šára. Dense stereomatching algorithm performance for view prediction and structure reconstruction. In *Proceedings of the Scandinavian Conference on Image Analysis*, volume LNCS 2749, pages 101–107, 2003.

[Kos93] A. Koschan. What is new in computational stereo since 1989: A survey on current stereo papers. Technical Report 93-22, Technische Universitaet Berlin, Fachbereich Informatik, August 1993.

[KŠ03] J. Kostková and R. Šára. Stratified dense matching for stereopsis in complex scenes. In *Proceedings of the British Machine Vision Conference*, volume 1, pages 339–348, 2003.

[Kv80] J. J. Koenderink and A. J. van Doorn. Photometric invariants related to solid shape. *Optica Acta*, 27(7):981–996, 1980.

[KZ01] V. Kolmogorov and R. Zabih. Computing visual correspondence with occlusions using graph cuts. In *Proceedings of the International Conference on Computer Vision*, pages 508–515, July 2001.

[KZ02] V. Kolmogorov and R. Zabih. Multi-camera scene reconstruction via graph cuts. In *Proceedings of European Conference on Computer Vision*, volume LNCS 2352, pages 82–96, 2002.

[KZ04] V. Kolmogorov and R. Zabih. What energy functions can be minimized via graph cuts? *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(2):147–159, February 2004.

[LZ99] C. Loop and Z. Zhang. Computing rectifying homographies for stereo vision. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 125–131, June 1999.

[Mar82] D. Marr. *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information.* W. H. Freeman and Company, 1982.

[MKZB01] S. Magda, D. J. Kriegman, T. Zickler, and P. N. Belhumeur. Beyond Lambert: reconstructing surfaces with arbitrary BRDFs. In *Proceedings of the International Conference on Computer Vision*, volume 2, pages 391–398, July 2001.

[MS05] K. Mikolajczyk and C. Schmid. A performance evaluation of local descriptors. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(10):1615–1630, 2005.

[MŠH04] M. Matoušek, R. Šára, and V. Hlaváč. Data-optimal rectification for fast and accurate stereovision. In *Proceedings of the Third International Conference on Image and Graphics*, pages 212–215, 2004.

[MTS+05] K. Mikolajczyk, T. Tuytelaars, C. Schmid, A. Zisserman, J. Matas, F. Schaffalitzky, T. Kadir, and L. V. Gool. A comparison of affine region detectors. *International Journal of Computer Vision*, 65(1-2):43–72, November 2005.

[NIK91] S. K. Nayar, K. Ikeuchi, and T. Kanade. Surface reflection: Physical and geometrical perspectives. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(7):611–634, July 1991.

[Ran06] *25 Years of RANSAC, Proceedings of a CVPR 2006 Workshop*. IEEE Computer Society Press, 2006.

[Ris78] J. Rissanen. Modeling by shortest data description. *Automatica*, 14:465–471, 1978.

[Šár02] R. Šára. Finding the largest unambiguous component of stereo matching. In *Proceedings of European Conference on Computer Vision*, volume LNCS 2352, pages 900–914, 2002.

[Šár06] R. Šára. Robust correspondence recognition for computer vision. In *Proceedings in Computational Statistics of 17th ERS-IASC Symposium*, pages 119–131. Physica-Verlag, 2006.

[Šár07] R. Šára. *The Principle of Stability Applied to Matching Problems in Computer Vision.* Habilitation thesis, Czech Technical University in Prague, Faculty of Electrical Engineering, 2007.

[SFS04] D. Schlesinger, B. Flach, and A. Shekhovtsov. A higher order MRF model for stereo reconstruction. In *Proceedings of the DAGM Symposium*, volume LNCS 3175, pages 440–446, 2004.

[ŠOS05] R. Šára, I. S. Okatani, and A. Sugimoto. Globally convergent range image registration by graph kernel algorithm. In *Proceedings of 5th International Conference on 3-D Digital Imaging and Modeling*, pages 377–384, 2005.

[SS02] D. Scharstein and R. Szeliski. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. *International Journal on Computer Vision*, 47(1-3):7–42, April-June 2002.

[SSŠ07] I. Shimizu, A. Sugimoto, and R. Šára. Graph-based range image registration combining geometric and photometric features. In *Proceedings of the Scandinavian Conference on Image Analysis*, volume LNCS 4522, pages 542–552, 2007.

[SWL98] W. Snyder, Z. Wan, and X. Li. Thermodynamic constraints on reflectance reciprocity and Kirchhoff's law. *Applied Optics*, 37(16):3464–3470, June 1998.

[Tor00] P. H. S. Torr. MLESAC: a new robust estimator with application to estimating image geometry. *Computer Vision and Image Understanding*, 78(1):138–156, April 2000.

[vH89] H. von Helmholtz. *Optique physiologique*, volume I, pages 231–232. Éditions Jaques Gabay, Sceaux, France, 1989.

[vNM44] J. von Neumann and O. Morgenstern. *Theory of Games and Economic Behaviour*. Princeton University Press, Princeton, 1944.

[YP84] A. L. Yuille and T. Poggio. A generalized ordering constraint for stereo correspondence. A.I. Memo 777, Artifical Intelligence Lab, Massachusetts Institute of Technology, May 1984.

[ZBK02] T. Zickler, P. N. Belhumeur, and D. J. Kriegman. Helmholtz stereopsis: Exploiting reciprocity for surface reconstruction. *International Journal on Computer Vision*, 49(2):215–227, 2002.

[ZW94] R. Zabih and J. Woodfill. Non-parametric local transforms for computing visual correspondence. In *Proceedings of European Conference on Computer Vision*, volume LNCS 801, pages 151–158, 1994.

Dr. Tech. Ing. Radim Šára

Affiliation: Center for Machine Perception, Department of Cybernetics, Faculty of Electrical Engineering (FEE), Czech Technical University in Prague (CTU), Technická 2, 166 27 Prague 6, Czech Republic. Phone: (+420) 224 357 203, fax: (+420) 224 357 385, E-mail: sara@cmp.felk.cvut.cz, URL: http://cmp.felk.cvut.cz/~sara/

Personal: born July 11, 1963, single

Appointments:

1997–present Research fellow at the Center for Machine Perception, Department of Cybernetics, FEE CTU in Prague.

1994–1996 Assistant professor, Department of Control Engineering FEE CTU in Prague

Membership: IEEE, Czechoslovak Pattern Recognition Society (IAPR member)

Education:

1987 Master degree in Electrical Engineering (Technical Cybernetics, Robotics) from FEE CTU Prague, Czechoslovakia. Thesis: *Classification and Image Description Using Relational Structures* (in Czech).

1994 PhD degree in Computer Science from Johannes Kepler Universität Linz, Austria. Thesis *Local Shading Analysis via Isophotes Properties.*

Academic Visits:

1992–1993 Johannes Kepler University Linz, Austria. Two-year student research fellowship.

1995–1997 GRASP Lab, School of Engineering and Applied Science, University of Pennsylvania. Two-year postdoctoral fellowship.

2000 Centre for Image Technology and Robotics, University of Auckland. Three-month visit.

2004 National Institute of Informatics, Tokyo, Japan

Selected Research Projects:

2007 Applied computer vision research for Volkswagen Aktiengesellschaft. Wolfsburg, Germany.

2007–2009 Partner in EU 6th framework STREP project *eTraining for Interpreting Images of Man Made Scenes* No. FP6-IST-027113.

2004–2008 Principal investigator of *Automatic 3D Virtual Model Builder from Photographs*, Project 1ET101210406 supported by the Grant Agency of the Czech Academy of Sciences.

2003–2005 Co-investigator of *Computer Aided Diagnosis of Diffuse Parenchyme Processes by Texture Analysis of Ultrasound Images.* In collaboration with 1st Medical Faculty of Charles University, Prague. Grant NO/7742–3 of the Ministry of Health of the Czech Republic.

2003–2005 Co-investigator of *Quantifying the Relation between Breathing Movements and the Function of the Spine*. In collaboration with P. Strnad, Faculty of Sports, Charles University, Prague. Grant NO/7735–3 of the Ministry of Health of the Czech Republic.

2001–2003 Principal investigator of *Physics Based Computer Vision*. Grant 102/01/1371 of the Grant Agency of the Czech Republic.

1999–2001 Co-investigator of *Texture Analysis of Sonographic Images for Endocrinopaties and Metabolic Diseases*. In collaboration with 1st Medical Faculty of Charles University, Prague. Grant NB 5472–3 of the Ministry of Health of the Czech Republic.

1995–1997 Principal investigator of *3-D Object Reconstruction from Intensity Images*. Grant 102/95/1378, Grant Agency of the Czech Republic.

Research Interests: Computer vision; geometric stereovision, reconstruction of 3-D models from uncertain visual information; photometric methods for 3-D shape reconstruction, shape from shading, local shading analysis, and photometric stereo; formalization of stability via digraph kernels.

Teaching: Course lectures: 3D Computer Vision (graduates), General Systems Theory (undergraduates), Technical Communication (undergraduates). Supervisor of MSc students, student supervisor in the PhD program of *Artificial Intelligence and Biocybernetrics*, two finished PhD students.

Service to scientific community: Co-organizer of *Int Conf on Computer Analysis of Images and Patterns*, Prague 1995. Co-organizer of *European Conf on Computer Vision*, and associated workshops, Prague 2004. Reviewer in *IEEE Trans on Pattern Recognition and Machine Intelligence*, *IEEE Trans on Medical Imaging*, *IEEE Trans on Image Processing*, *Int J on Computer Vision*, *Machine Vision Applications*, *Image Vision and Computing*, *Machine Graphics and Vision*, *J of Optical Society of America (Ser. A)*, *Pattern Recognition*, Programme Committee Member in *IEEE Computer Soc Conf on Computer Vision and Pattern Recognition 2005–2007*, *IEEE Int Conf on Computer Vision 2005–2007*, *European Conf on Computer Vision 2006*, *Asian Conf on Computer Vision 2007*, and others.

Cited work (SCI w/o direct and indirect self-citations): [10] 22×, [8] 12×, [12] 9×, [13] 8×, [11] 7×, [15] 6×, [14] 4×, [9] 2×, [6] 1×, unlisted work 21×.

Selected Publications Characterizing My Work Well:

[1] M. Bujňák and R. Šára. A robust graph-based method for the general correspondence problem demonstrated on image stitching. In *Proceedings of 11th IEEE International Conference on Computer Vision*, October 2007.

[2] J. Kostlivá, J. Čech, and R. Šára. Feasibility boundary in dense and semi-dense stereo matching. In *BenCOS 2007: CVPR Workshop Towards Benchmarking Automated Calibration, Orientation and Surface Reconstruction from Images*, IEEE Computer Society, June 2007. Best Paper Award.

[3] R. Šára. Robust correspondence recognition for computer vision. In *COMPSTAT 2006: Proceedings in Computational Statistics of 17th ERS-IASC Symposium*, pp. 119–131. Physica-Verlag, August/September 2006.

[4] R. Šára and M. Matoušek. FAIR: Towards a new feature for affinely-invariant recognition. In *Proceedings of the International Conference on Pattern Recognition*, vol. 2, pp. 412–416. IEEE Computer Society Press, August 2006.

[5] V. Smutný and R. Šára. Method and a system for measuring joint trajectory, particularly mandibular joint trajectory. Patent 296088, Úřad průmyslového vlastnictví Praha, November 2005.

[6] R. Šára, I. S. Okatani, and A. Sugimoto. Globally convergent range image registration by graph kernel algorithm. In *3DIM 2005: Proceedings of 5th International Conference on 3-D Digital Imaging and Modeling*, pp. 377–384, IEEE Computer Society, June 2005.

[7] Z. Jankó, O. Drbohlav, and R. Šára. Radiometric calibration of a Helmholtz stereo rig. In *Proc IEEE Int Conf on Computer Vision*, vol. 1, pp. 166–171, June/July 2004.

[8] D. Smutek, R. Šára, P. Sucharda, T. Tjahjadi, and M. Švec. Image texture analysis of sonograms in chronic inflammations of thyroid gland. *Ultrasound in Medicine and Biology*, 29(11):1531–1543, November 2003.

[9] D. Smutek, R. Šára, and P. Sucharda. Relation between quantitative description of ultrasonographic image and clinical and laboratory findings in lymphocytic thyroiditis. *Endocrine Regulations*, 37(3):181–187, September 2003.

[10] R. C. Gur, R. Šára, M. Hagendoorn, O. Marom, P. Hughett, L. Macy, T. Turner, R. Bajcsy, A. Posner, and R. E. Gur. A method for obtaining 3-dimensional facial expressions and its standardization for use in neurocognitive studies. *Journal of Neuroscience Methods*, 115(2):137–143, April 2002.

[11] O. Drbohlav and R. Šára. Specularities reduce ambiguity of uncalibrated photometric stereo. In *Proc 7th European Conf on Computer Vision*, Springer LNCS 2351:46–60, May 2002.

[12] R. Šára. Finding the largest unambiguous component of stereo matching. In *Proc 7th European Conf on Computer Vision*, Springer LNCS 2352:900–914, May 2002.

[13] O. Drbohlav and R. Šára. Unambiguous determination of shape from photometric stereo with unknown light sources. In *Proc 8th Int Conf on Computer Vision*, vol. 1, pp. 581–586, July 2001.

[14] R. Šára and R. Bajcsy. Fish-scales: Representing fuzzy manifolds. In *Proc Int Conf on Computer Vision*, pp. 811–817, January 1998.

[15] R. Šára and R. Bajcsy. On occluding contour artifacts in stereo vision. In *Proc Int Conf Computer Vision and Pattern Recognition*, pp. 852–857, June 1997.

In Prague, February 29, 2008 Radim Šára