

Image Comparison on the Base of a Combinatorial Matching Algorithm

Benjamin Drayer

Department of Computer Science, University of Freiburg

Abstract. In this paper we compare images based on the constellation of their interest points. The fundamental technique for this comparison is our matching algorithm, that is capable to model miss- and multi-matches, while enforcing one-to-one matches. We associate an energy function for the possible matchings. In order to find the matching with the lowest energy, we reformulate this energy function as Markov Random Field and determine the matching with the lowest energy by an efficient minimization strategy. In the experiments, we compare our algorithm against the normalized cross correlation and a naive forth-and-back best neighbor match algorithm.¹

1 Introduction

The comparison of complex structures plays an important role in biological and medical research. In many cases the problem of how similar two structures are is posed on a scale where we have one to one correspondences such as eye to eye, leg to leg or mouth to mouth. A good strategy is to first perform a registration (e.g. an elastic registration such as [2]) and then measure the similarity of the pixel intensities e.g. with differences or with the normalized cross correlation (NCC). Difficulties arise, when there are no one-to-one correspondences guaranteed as shown in Figure 1.

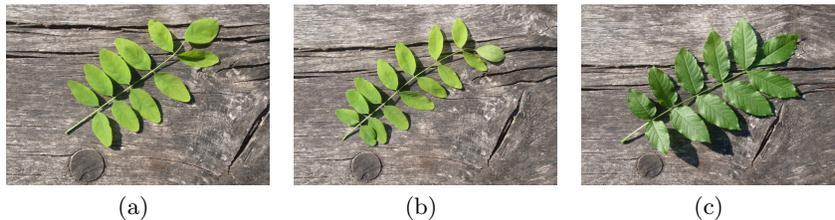


Fig. 1. While (a) and (b) match better according to the shape of the leaves, (a) and (c) are considered more similar regarding the number and arrangement of leaves. In this paper we define a matching based similarity measure that allows to consider both criteria

¹ recommended for submission to YRF2011 by Junior-Prof. Dr. Olaf Ronneberger

Obviously every intensity based similarity measure fails, when one leaf corresponds to no or several leaves. This problem also arises in microscope imaging, where we can register and compare different recordings on a rough scale to a certain degree. But when we go in detail, in our case the cell level, the slight differences between the same organs in different individuals become noticeable. One cell may correspond to no cell, to one cell or to many cells and vice versa.

To solve this problem, we apply a similarity measure based on correspondences between interest points of the images. Our main contribution is the matching algorithm that establishes these correspondences. An intuitive way for such a matching is a forth- and back-matching of the best neighbors. The disadvantage of this naive method is, that it cannot favor the one-to-one correspondences. As a result, we get un-proportional many multi-matches and the structure of the image is lost. Whereas our matching algorithm, matches not one key point to its closest neighbor, but it matches a local arrangement so that the one-to-one correspondences are favored and local constellations are properly taken into account.

2 Interest Points and Features

In our application, the substructures that we want to match are roundish. We detect them with a Laplacian of Gaussian as done in [4]. The Laplacian of Gaussian gets a high response not only on round structures but also on elongated structures such as edges and ridges. As in [4], [1], we remove the unstable key points caused by elongated structures and we reject the interest points in the background with an Otsu thresholding.

Considering 2D images, we describe the interest points with local, normalized color histograms. When we deal with 3D images, we compute rotation invariant features based on the spherical harmonics [3], where the spherical harmonics are computed on multiple spheres at different scales in order to better describe the volume around the interest point.

3 Matching Algorithm

With each matching f , we associate an energy in such a way that the one-to-one correspondences are favored. Minimizing the energy function in a brute force manner is too expensive and since the energy function is non-convex, a gradient descent is not an option. Therefore we reformulate the energy as a markov random field and use a combinatorial approach.

For the matching, we consider the sets \mathcal{A} and \mathcal{B} , each containing vectors with geometrical and appearance information of the respective interest points. With $d(\mathbf{a}, \mathbf{b})$ we denote the distance between the interest points in the feature space, where $\mathbf{a} \in \mathcal{A}$ and $\mathbf{b} \in \mathcal{B}$. The distance is the l_2 -norm, normalized by the average distance of the best neighbor (in \mathcal{A}) of each interest point in \mathcal{A} .

With f we denote the matching relation from \mathcal{A} to \mathcal{B} . Equivalently, we can write that as

$$f : \mathcal{A} \rightarrow \mathcal{P}(\mathcal{B}), \quad (1)$$

where \mathcal{P} denotes the power set.

In order to judge the quality of a certain matching, we compute an energy function $E : f \rightarrow \mathbb{R}_+$. This energy function consist of the cost for the matchings from \mathcal{A} to \mathcal{B} and the costs for the not matched points in \mathcal{B} . With $E_f(\mathbf{a})$ we denote the cost that is contributed by \mathbf{a} under the mapping f :

$$E_f(\mathbf{a}) = \begin{cases} p_\emptyset & f(\mathbf{a}) = \emptyset \quad (\text{miss match}) \\ d(\mathbf{a}, \mathbf{b}) & f(\mathbf{a}) = \{\mathbf{b}\} \quad (\text{single match}) \\ p_m + \sum_{\mathbf{b}_i \in f(\mathbf{a})} d(\mathbf{a}, \mathbf{b}_i) & |f(\mathbf{a})| > 1 \quad (\text{multi match}) \end{cases}, \quad (2)$$

where p_\emptyset and p_m are constants that penalize miss- and multi-matches. In our experiments, we choose the average of the distance for p_\emptyset and $p_m = 10$.

So far this energy favors a one-way best neighbor match. We introduce another energy term for the not matched elements from \mathcal{B} . This term $E_f(\mathbf{b})$ forces the one-to-one correspondences and therefore it also preserves the local structure. Let \mathbf{a}_1 and \mathbf{a}_2 be the two closest elements to \mathbf{b} . Then we define a miss-match penalty function:

$$p(\mathbf{b}) = \min(\lambda d(\mathbf{a}_1 - \mathbf{b}) + (1 - \lambda)d(\mathbf{a}_2 - \mathbf{b}), p_\emptyset), \quad (3)$$

with $\lambda \in [0, 1]$. For our experiments, we choose $\lambda = 0.5$.

For the assignment of the miss-match cost, we need information of the inverse matching relation. The inverse matching relation f_{inv} of \mathbf{b} is:

$$f_{\text{inv}}(\mathbf{b}) = \{\mathbf{a} \in \mathcal{A} | \mathbf{b} \in f(\mathbf{a})\}. \quad (4)$$

The energy term $E_f(\mathbf{b})$ becomes:

$$E(\mathbf{b}) = \begin{cases} p(\mathbf{b}) & f_{\text{inv}}(\mathbf{b}) = \emptyset \\ 0 & \text{else} \end{cases}. \quad (5)$$

Altogether the energy function of mapping f is:

$$E(f) = \sum_{\mathbf{a} \in \mathcal{A}} E_f(\mathbf{a}) + \sum_{\mathbf{b} \in \mathcal{B}} E_f(\mathbf{b}). \quad (6)$$

3.1 Reformulation as Markov Random Field

The brute force solution for this problem is too expensive (exponential runtime), even if we restrict the amount of mappings in such a way that an \mathbf{a} can only be mapped to an arbitrary subset of its k closest \mathbf{b} . Therefore, we reformulate the problem as a Markov Random Field of the form:

$$E = \sum_{\mathbf{a} \in \mathcal{V}} V_{\mathbf{a}}(l_{\mathbf{a}}) + \sum_{(\mathbf{a}_i, \mathbf{a}_j) \in \mathcal{E}} V_{\mathbf{a}_i \mathbf{a}_j}(l_{\mathbf{a}_i} l_{\mathbf{a}_j}) \quad (7)$$

and can apply sophisticated algorithms such as the *max-sum* solver [5] to minimize the energy in a reasonable time.

The set of nodes \mathcal{V} is \mathcal{A} . The set of labels for each \mathbf{a} are the 2^k possible subsets of its k closest \mathbf{b} . In our implementation, we choose $k = 5$. The set of edges \mathcal{E} consists of pairs of nodes $(\mathbf{a}_1, \mathbf{a}_2)$ that have at least one common \mathbf{b} in their sets of k closest \mathbf{b} , otherwise they don't form an edge.

The unary potential becomes:

$$V_{\mathbf{a}}(l_{\mathbf{a}}) = \begin{cases} p_{\emptyset} & l_{\mathbf{a}} = \emptyset \quad (\text{miss match}) \\ 0 & |l_{\mathbf{a}}| = 1 \quad (\text{single match}) \\ p_m & |l_{\mathbf{a}}| > 1 \quad (\text{multi match}) \end{cases}. \quad (8)$$

The pairwise potential becomes:

$$V_{\mathbf{a}_i, \mathbf{a}_j}(l_{\mathbf{a}_i}, l_{\mathbf{a}_j}) = \sum_{\mathbf{b} \in \{f(\mathbf{a}_i) \cap f(\mathbf{a}_j)\}} \frac{E_{\mathbf{a}_i, \mathbf{a}_j, \mathbf{b}}(l_{\mathbf{a}_i}, l_{\mathbf{a}_j})}{\binom{|f_{\text{inv}}(\mathbf{b})|}{2}} \quad (9)$$

, where

$$E_{\mathbf{a}_i, \mathbf{a}_j, \mathbf{b}}(l_{\mathbf{a}_i}, l_{\mathbf{a}_j}) = \begin{cases} p(\mathbf{b}) & \mathbf{b} \notin l_{\mathbf{a}_i} \wedge \mathbf{b} \notin l_{\mathbf{a}_j} \quad (\text{miss match}) \\ d(\mathbf{a}_i, \mathbf{b}) & \mathbf{b} \in l_{\mathbf{a}_i} \wedge \mathbf{b} \notin l_{\mathbf{a}_j} \quad (\text{single match}) \\ d(\mathbf{a}_j, \mathbf{b}) & \mathbf{b} \notin l_{\mathbf{a}_i} \wedge \mathbf{b} \in l_{\mathbf{a}_j} \quad (\text{single match}) \\ d(\mathbf{a}_i, \mathbf{b}) + d(\mathbf{a}_j, \mathbf{b}) & \mathbf{b} \in l_{\mathbf{a}_i} \wedge \mathbf{b} \in l_{\mathbf{a}_j} \quad (\text{multi match}) \end{cases}. \quad (10)$$

We normalize by $\binom{|f_{\text{inv}}(\mathbf{b})|}{2}$, since this is the number of pairs $(\mathbf{a}_1, \mathbf{a}_2)$ that have \mathbf{b} in common.

4 Experiments

We show the performance of our algorithm with two experiments. First, we compute correspondences for different constellations of fruits and second, we apply the matching algorithm on microscopic data of the zebra fish embryo. The correspondences are the base for the similarity measure, where only one-to-one correspondences are taken into account (multi-matches inconsistencies are reduced to their best edge). The similarity measure is:

$$s(\mathcal{A}, \mathcal{B}) = \frac{2}{|\mathcal{A}| + |\mathcal{B}|} \sum_{(\mathbf{a}, \mathbf{b}) \in \mathcal{S}} \frac{1}{1 + d(\mathbf{a}, \mathbf{b})}, \quad (11)$$

where \mathcal{S} denotes the set of one-to-one correspondences.

In Figure 2 we compare our algorithm against the NCC. In the first row of this Figure, we try to match a mixture of apples and kiwis, where the position and the amount of the objects vary. In the second row, we almost keep the position of the fruits, do not vary the amount, but we change the kiwis to apples

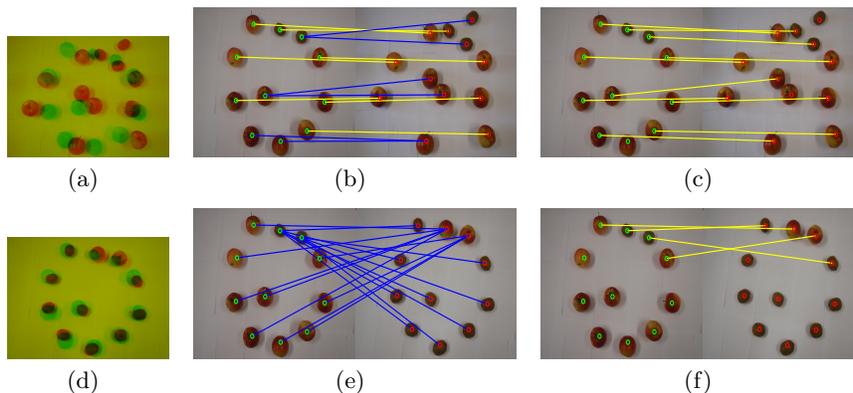


Fig. 2. (a) overlay of the images, similarity=0.2969 (NCC) (b) all edges of multi-match. (c) confident edges of multi-match, similarity=0.836 (d) overlay of the images, similarity=0.7401 (NCC) (e) all edges of multi-match. (f) confident edges of multi-match, similarity=0.331

and vice versa. For the NCC, the case in the second row is more similar, whereas with our matching algorithm we correctly determine that the images in the first row are more similar.

The second part of this experiment (Figure 3) shows the superiority of our multi-matching algorithm over the best neighbor match. By comparing the two sets of apples, arranged as bows we see advantageous effects of matching a local neighborhood, especially the preserving of the local structure. The histogram of the matches in Figure 3(d) shows the strong favoritism of the one-to-one correspondences of our approach.

In Figure 4, we show the performance of our matching algorithm on microscopic recordings of the zebra fish embryo. The recordings are taken on the cell level and a dataset measures $800 \times 500 \times 500$ voxel. Based on our matching algorithm, we can correctly classify a large database of gene expressions, but as this is a part of a joint unpublished project, only one example can be shown here.

5 Conclusion

In this paper we presented a matching algorithm for a correspondence based similarity measure. We showed the advantages of our comparison method over the normalized cross correlation and the best neighbor match algorithm. Furthermore, we got promising results when we applied the algorithm on microscopic recordings of the zebra fish.

References

1. Allaire, S., Kim, J.J., Breen, S.L., Jaffray, D.A., Pekar, V.: Full orientation invariance and improved feature selectivity of 3D SIFT with application to medical image

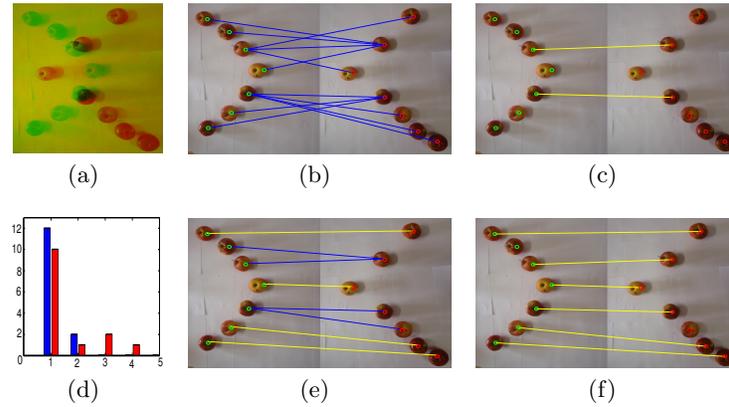


Fig. 3. (a) overlay of the images. (b) all edges of best match. (c) confident edges of best match, similarity=0.412. (d) distribution of matches for the best match (red) and multi-match (blue) algorithm. (e) all edges of multi-match. (f) confident edges of multi-match, similarity=0.748. The blue edges indicate multiple correspondences, the yellow edges indicate one-to-one correspondences

- analysis. Computer Vision and Pattern Recognition Workshops (Jan 2008)
2. Glocker, B., Komodakis, N., Tziritas, G., Navab, N., Paragios, N.: Dense image registration through mrfs and efficient linear programming. *Medical Image Analysis* 12(6), 731–741 (2008)
 3. Kazhdan, M., Funkhouser, T., Rusinkiewicz, S.: Rotation invariant spherical harmonic representation of 3d shape descriptors. In: *Proceedings of the 2003 Eurographics/ACM SIGGRAPH symposium on Geometry processing*. pp. 156–164. SGP '03 (2003)
 4. Lowe, D.G.: Distinctive image features from scale-invariant keypoints. *Int. J. Comput. Vision* 60(2), 91–110 (2004)
 5. Werner, T.: A linear programming approach to max-sum problem: A review. *IEEE Trans. Pattern Anal. Mach. Intell.* 29, 1165–1179 (July 2007)

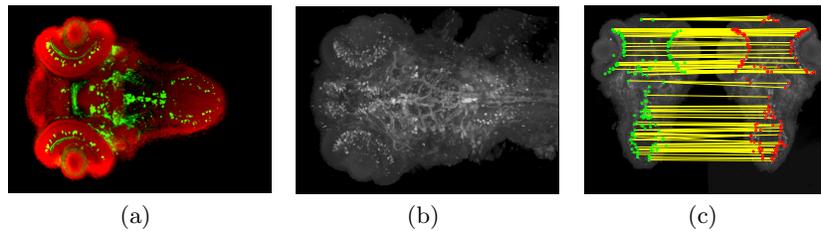


Fig. 4. (a) cells (red) and gene expression (green), (b) maximum intensity projections of the gene expression (c) matching of the same gene expression (of different individuals), green and red circles are the found interest points, yellow lines indicate the correspondences. Due to better visibility this is a maximum intensity projection in z -direction from 220 to 250 μm