Volume xx (200y), Number z, pp. 1–7

Partial Shape Matching using Transformation Parameter Similarity - Additional Material

paper1066

1. An Example for Local Contour Descriptors

As mentioned in Section 4 of the paper, the local descriptors used for our method are interchangeable. In this Section we describe one possible choice of local descriptors for shapes given as contour segments, based on the well-known Shape Contexts [BMP02].

To define the placement of the descriptors, we sample the shape contours with a discrete set of points. Sampling is performed depending on local contour curvature, creating more points in regions of high curvature [CFH*09]. In order to prevent straight lines from being undersampled, a fixed minimum distance (about one percent of the major image dimension) to neighboring points along a contour is always respected. The orientation of each descriptor is aligned with the local tangent at the contour sample. When comparing two descriptors, both possible orientations are tested. To determine the scale of a descriptor at a contour sample, we compute the arclength a of a contour interval centered at the contour sample having a small constant curvature integral c. The scale is then proportional to $\frac{a}{c}$. We found that this value is a robust measure of the contour line scale.

Now that the position, scale and angle of all descriptors is defined, we proceed to compute the descriptor signature, i.e. a vector that describes the geometry in the local neighborhood of the descriptor. The Shape Context descriptor [BMP02] provides such a signature and is known to work well on shapes given as point sets. We describe a slightly modified version that is more robust to background clutter and slight variations in the position of contour lines. The local area around a point is discretized into N bins, each bin holding the weighted average of all contour tangents inside a fixed distance from the bin center and a weight describing the distance of the bin center to the closest contour (see Figure 1). The distance weight for bin k of shape context i is defined as

$$d_i^k = max(0, 1 - \frac{d_c}{d_{max}\ rad_i}) \tag{1}$$

where d_c is the distance of the bin center to the closest contour line, rad_i is the radius of the shape context and d_{max} is a maximum distance defined by the user. This function models a linear falloff in a band around the contour lines, improving the robustness of the shape context to slight variations in the positions of contour lines. The constant d_{max} controls the width of the falloff, we determined empirically that 0.3 is a good value. The tangent for a bin k is a weighted average over the tangents of all neighboring contours inside the radius d_{max}, with weights based on the distance to the bin center.

Before comparing two shape contexts, we align their frames of reference. The matching confidence for two shape descriptors is then

$$f_{ij} = \frac{\sum_k d_i^k d_j^k (1 - ||g_i^k - g_j^k||^2)}{\sum_k d_i^k d_j^k}, \tag{2}$$

where g_i^k is the average tangent in bin k of the shape context i. Note that this confidence measure is different from the original measure proposed in [BMP02]. Our measure reduces the effect of background clutter by rewarding matched bins, rather than penalizing unmatched bins. The first term $d_i^k d_j^k$ inside the sum measures the agreement of the contour line positions in both shape contexts. Note that this term is only nonzero if the contour lines are within a threshold of $2 \cdot d_{max}$. For contours that are present in one shape context, but not in the other one, the

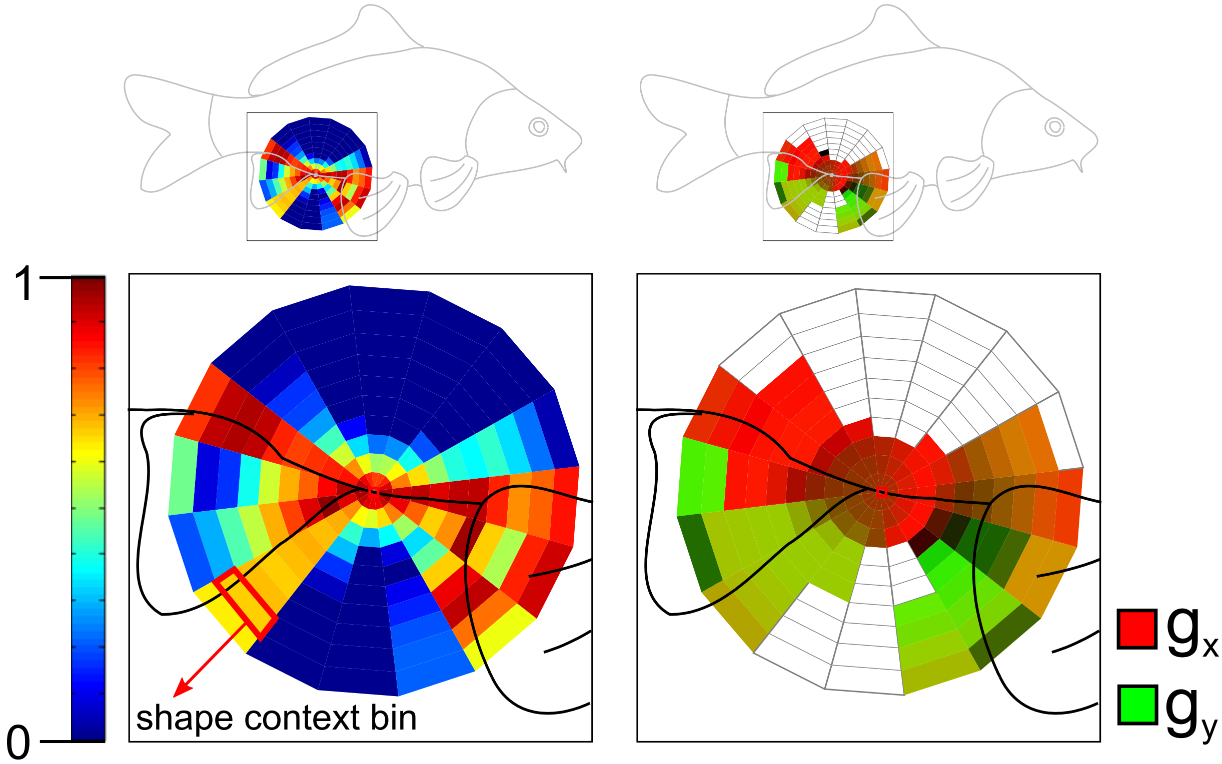

Figure 1: *We store the distance to the contours and the average contour normal in each bin of the shape contexts. The distance is shown on the left, the right side shows the average tangent per bin.*

first term is zero and the sum remains unchanged. The second term measures the agreement of the contour line orientations.

All pairs of descriptors with high matching confidence can be used to form the first-order correspondences (p, τ) as described in the paper, where τ are the parameters of the transformation from the local reference frame of descriptor one (including position, orientation and scale) to the local reference frame of descriptor two.

If the contours are noisy the orientations of the shape context and consequently the rotation parameter of first-order correspondences become less reliable. Since the transformation of a correspondence is defined as $T(t_x, t_y, s, r) = T^L(t_x, t_y) * T^S(s) * T^R(r)$ (see Section 3 in the paper), this also affects the translation parameters by an amount proportional to the distance between the correspondence and the coordinate origin. The correspondence refinement step alleviates this problem. To further improve the stability of the translation parameters in the presence of unreliable descriptor orientations, we could dynamically change the coordinate origin when computing the gradient of the translation parameters. When computing the linkage weights a_{ij} (Equation 6 in the paper), we can re-compute the translation parameters of both correspondences (q_i, τ_i^*) and (q_j, τ_j^*) before computing the difference quotients by moving the coordinate origin to the query point q_i of the first correspondence. The new translation parameters for both correspondences (q_*, t_*) with $* = i$ or j are then $(t_x, t_y) = (p_* - q_i) \; - \; T^S(s_*) * T^R(r_*) * (q_* - q_i)$, where $p_* = T(\tau_*) * q_*$ is the target position of the correspondence. Note that this results in an asymmetric linkage weight matrix. To re-symmetrize it, we can take the minimum of a_{ij} and a_{ji} for each entry of the matrix.

2. Transformation Space Visualizations

In Figures 2 – 5 we show visualizations of the transformation space of several example images. In the top row of each figure, refined first-order correspondences for the scene on the left are shown in shades of grey, where darker correspondences have higher confidence. In the bottom row, we show the N best matches found by our region matching step, the subsets of correspondences defining each matched regions are shown in different colors. N is picked by the user after the regions have been found. The goal during the region matching step of our method is to find subsets of refined correspondences that form transformation functions with low Jacobian magnitude as shown in the bottom row. The center of the figure shows plots of the individual transformation parameters of all correspondences, while the rightmost plot shows the correspondences in three of the four transformations space dimensions. Note how the first-order correspondences form separable clusters in transformation space (rightmost plot). Additionally, each cluster forms a transformation function with low Jacobian magnitude (see the four central plots). The caption of each figure additionally provides the inlier ratio of the first-order correspondences and the inlier ratio weighted by the confidence of the correspondences.

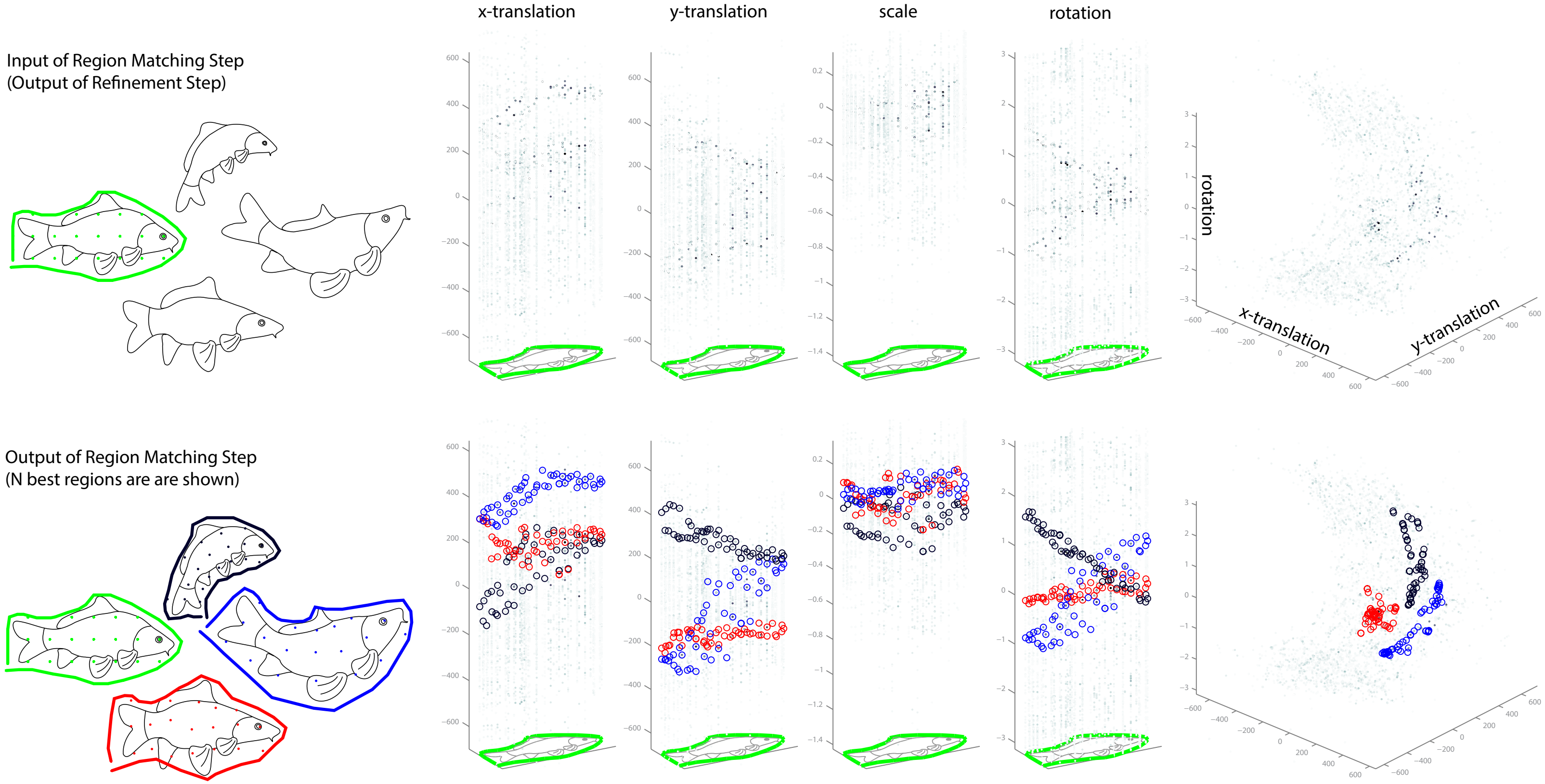

Figure 2: *A simple example where the shape of the clusters in transform space is clearly recognizable. There are a total of* 5759 *refined first-order correspondences and the inlier ratio is* 2.62% *(weighted:* 12.99%*).*

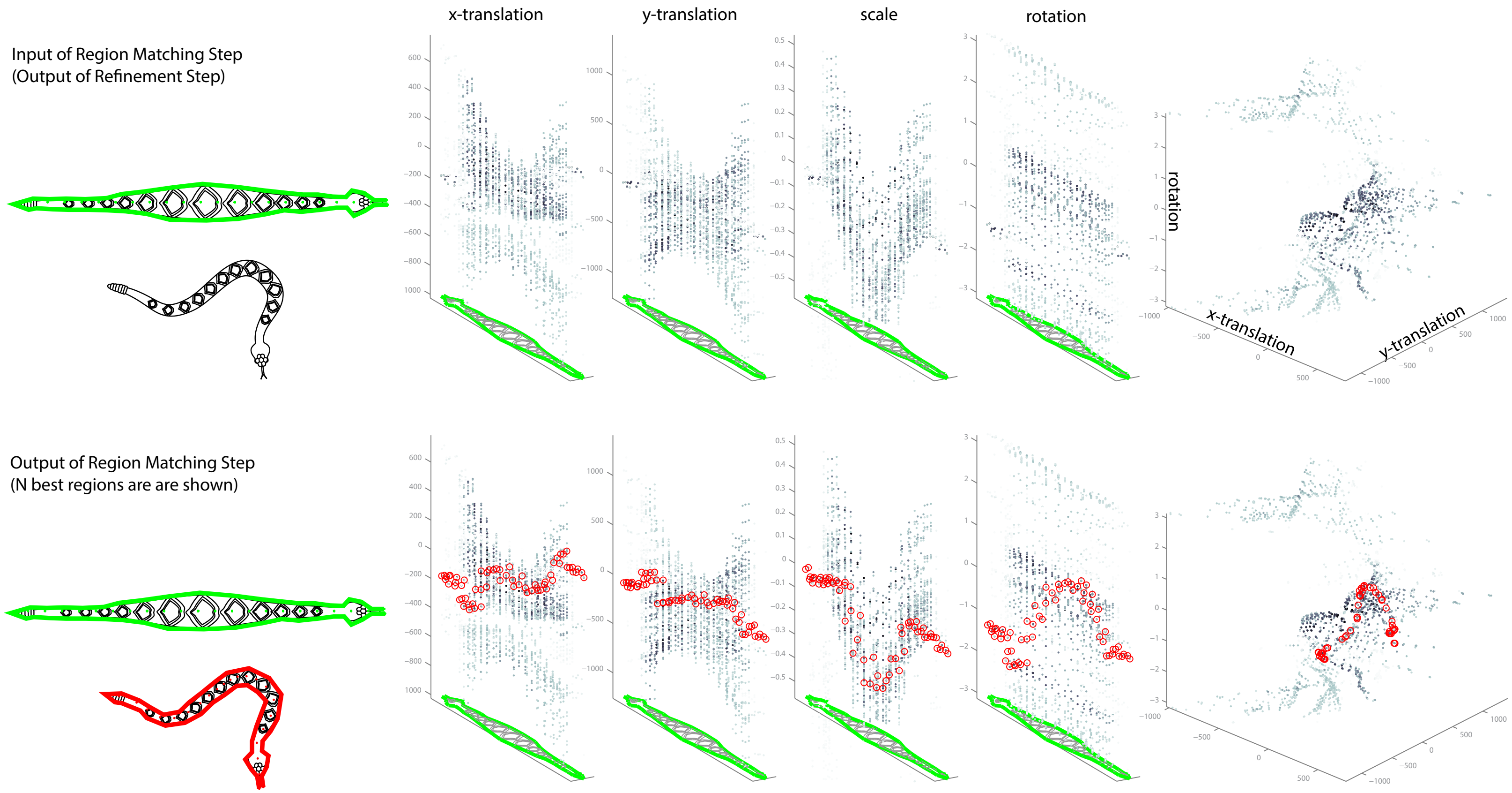

Figure 3: *The snakes image shown in the paper. Note how the repeating pattern of the skin creates a lot of high-confidence outliers. There are a total of* 2551 *refined first-order correspondences and the inlier ratio is* 2.47% *(weighted:* 4.18%*).*

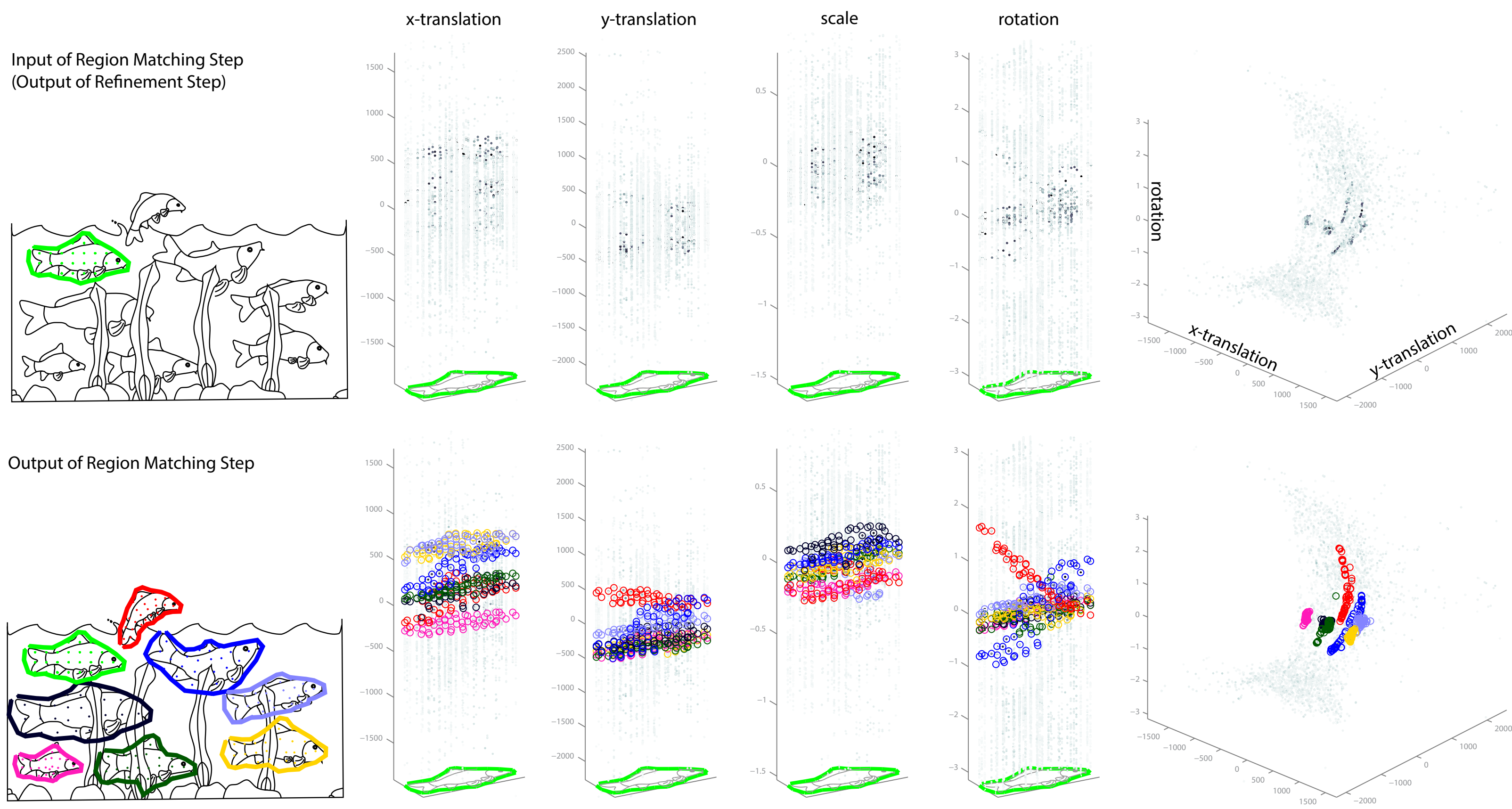

Figure 4: *The fish image shown in the paper. The background clutter produces a large amount of low-confidence correspondences. Note that we cannot simply threshold the correspondences, as matched region contain both low- and high confidence correspondences. There are a total of* 7810 *refined first-order correspondences and the inlier ratio is* 4.34% *(weighted:* 19.65%*).*

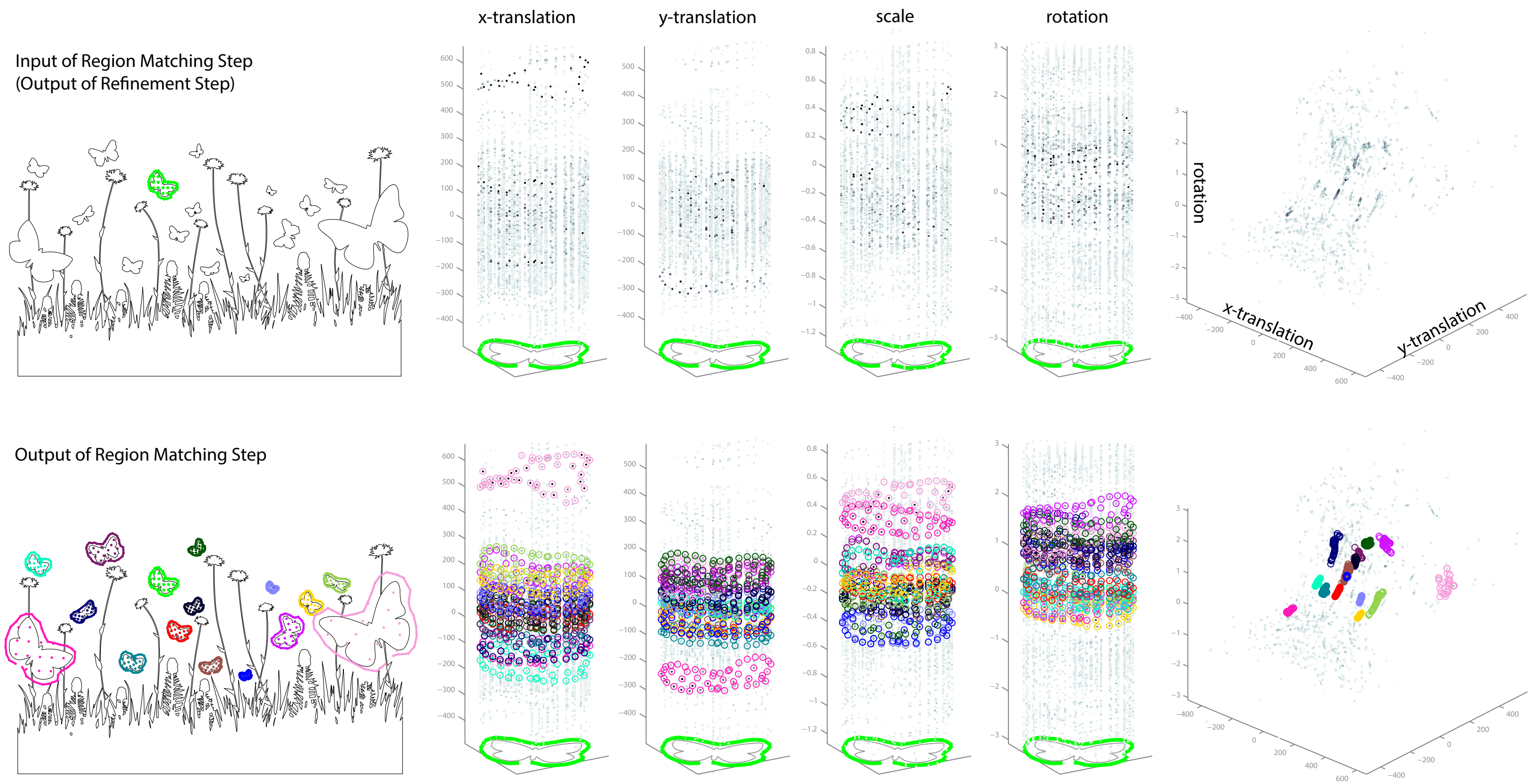

Figure 5: *The butterflies image shown in the paper. Like in the fish image, the background clutter produces a large amount of low-confidence correspondences. There are a total of* 5668 *refined first-order correspondences and the inlier ratio is* 12.35% *(weighted:* 29.69%*).*

References

[BMP02] BELONGIE S., MALIK J., PUZICHA J.: Shape matching and object recognition using shape contexts. *IEEE PAMI 24* (2002), 509–522. 1

[CFH*09] CUI M., FEMIANI J., HU J., WONKA P., RAZDAN A.: Curve matching for open 2d curves. *Pattern Recogn. Lett. 30*, 1 (2009), 1–10. 1