

Effective Visualization of Sparse Image-to-Image Correspondences

Supplemental Material

C. Andujar¹, A. Chica¹, M. Comino¹

¹VirVIG, Computer Science Department, Universitat Politecnica de Catalunya, Jordi Girona 1-3, Barcelona, Spain

1. Additional results

Figures 1 and 2 compare different clustering strategies with a baseline approach. The best strategy depends on the user's interest on outlier vs inlier matches. The *min* strategy tends to over-aggregate consistent matches, resulting in a few large aggregated segments showing the overall match directions, plus other segments representing isolated matches (outliers). The *min* strategy thus shows most outliers, but due to over-aggregation it might hide wrong matches not deviating much from inlier match locations. The *max* strategy only groups highly-consistent matches, and thus the main match directions are represented by multiple aggregated segments. This in turn causes some outlier matches to be grouped onto the closest cluster, thus reducing the number of outlier matches that are shown as aggregated segments. The *avg* strategy is a trade-off between these two and thus the default option in our tool. The baseline approach allows users to spot some outlier directions, but due to occlusion many outliers are hidden and those that can be spotted cannot be traced visually end-to-end. Our output provides less-cluttered images and thus the benefits in terms of visual clarity overcome the loss of information.

Figures 3 and 4 show the benefits of hierarchical clustering. In the context of an interactive application, aggregated segments can be selected to show the underlying individual matches.

Finally, Figures 5 and 5 demonstrate the usefulness of our approach with challenging image pairs. Large scale variations between images reduce the likelihood of getting roughly parallel segments. When combined with image layout optimization, output drawings have clearly distinguishable paths.

2. Layout optimization details

We need to find the optimal layout for two images A and B , with sizes $w_a \times h_a$ and $w_b \times h_b$, resp. Between both images we have N matches with endpoints $a_i = (x_i, y_i) \in A$ and $b_i = (x'_i, y'_i) \in B$. Given that we will only allow for translations, we may consider that image A is fixed and image B is translated by a vector $t = (t_x, t_y)$. The optimization criterion will be the distance between matches, but as we do not want to have overlapping, we will restrict their relative positions by adding constraints. The error function without constraints will be:

$$E(t) = \sum_i \|a_i - b_i - t\|^2 = \sum_i (a_i - b_i - t)^T (a_i - b_i - t)$$

$$E(t) = \sum_i (a_i - b_i)^T (a_i - b_i) + Nt^T t - 2 \sum_i (a_i - b_i)^T t$$

We have to find a vector t that minimizes $E(t)$, but the first term does not depend on t , so:

$$\min_t E(t) = \min_t \left[Nt^T t - 2 \sum_i (a_i - b_i)^T t \right]$$

Now, if we want no overlap between the two images, we need to add constraints. We allow for two types of layouts, either side-by-side or one image on top of the other. From this point forward we will assume that the images are laid out horizontally. In this case, we will want to have $t = (s, t_y)$, with s a fixed value. When $s = w_a$ image B will be to the right of image A , while when $s = -w_b$ image B will be to the left of A . Substituting in the previous equations and eliminating from the minimization those terms that do not depend on t_y :

$$\min_t E(t) = \min_t \left[N(t_x t_x + t_y t_y) - 2 \sum_i t_x (x_i - x'_i) + t_y (y_i - y'_i) \right]$$

$$\min_t E(t) = \min_t \left[N t_y t_y - 2 \sum_i t_y (y_i - y'_i) \right]$$

Now we can derive and equal zero:

$$\frac{\partial E(t)}{\partial t} = 2N t_y - 2 \sum_i (y_i - y'_i) = 0$$

$$t_y = \frac{1}{N} \sum_i (y_i - y'_i)$$

Therefore, the optimal t_y is the average of the differences of the Y components of the matches. This is always true regardless of the

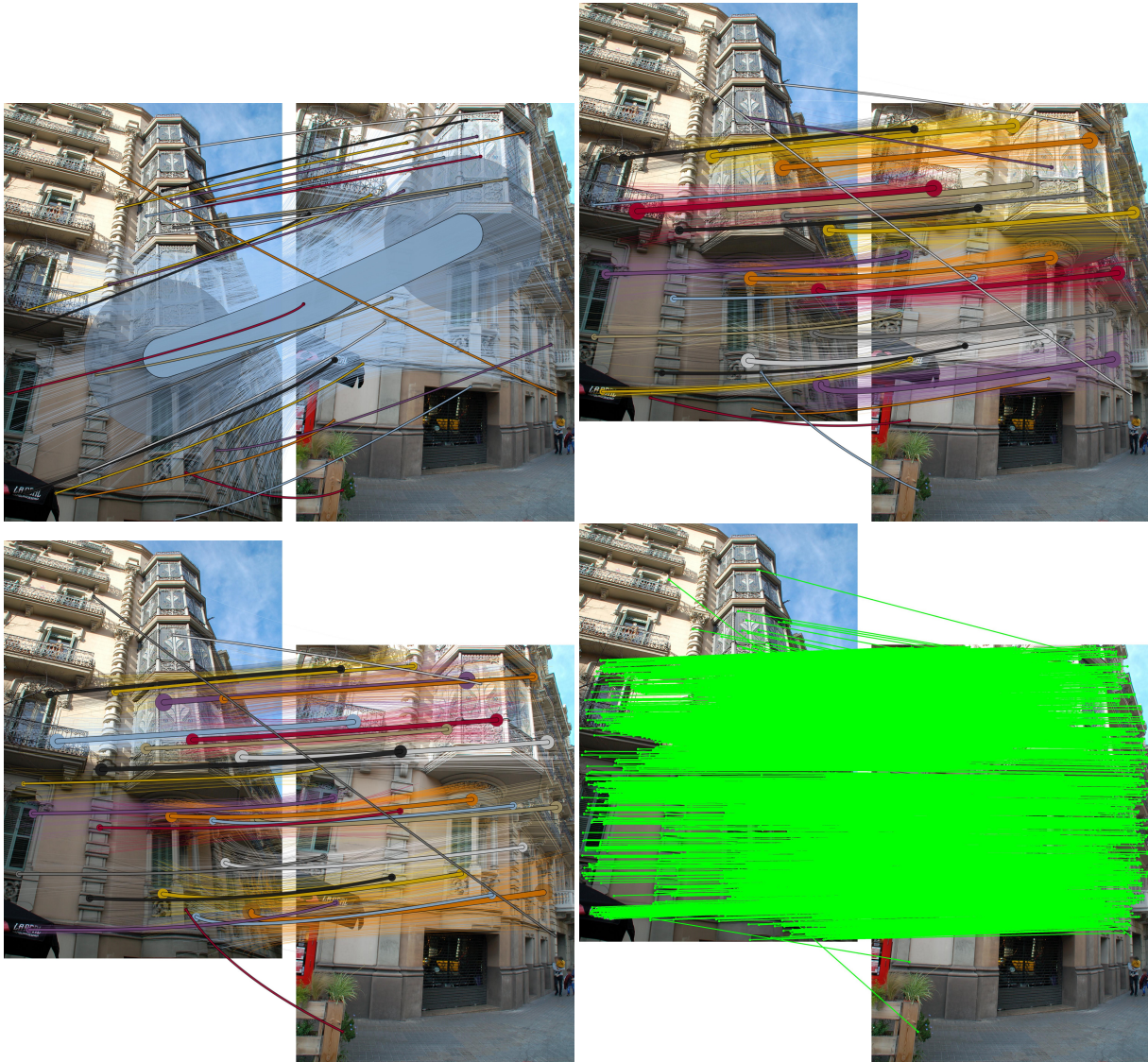


Figure 1: Aggregated segments using *min*, *avg* and *max* methods on a facade, compared with a baseline showing all matches. The resulting segments affect the optimized image layout, due to the varying aggregation of outlier directions.

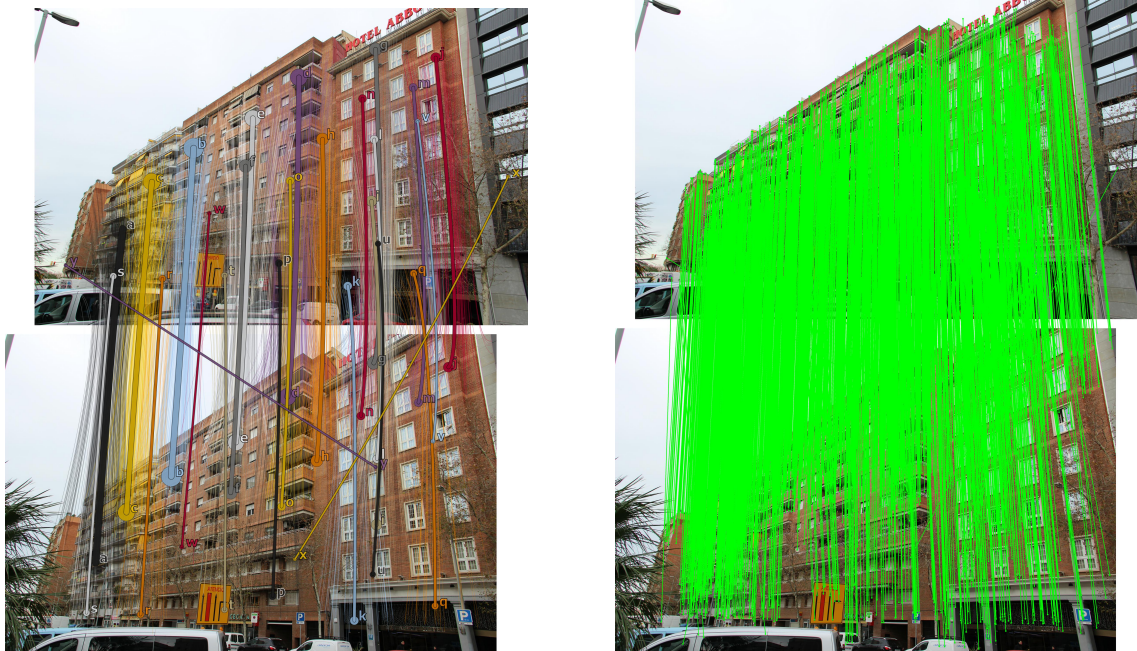
chosen s value, but we will have different error values for each s chosen. It is necessary to calculate $E(t)$ for $s = w_a$ and for $s = -w_b$ independently, and choose the option that results in the minimum error.

When images are arranged vertically the optimal t_x is also the average of the X components of the matches. In order to consider all the options, we should compute the optimal positions for the two horizontal layout cases and the two vertical ones separately, and then choose the one with the least associated error.



(a) Aggregated segments using *min*. We can observe outliers that stand out such as *h, e, i, w, y* and others more subtle such as *b, k, m, o, p, q, r, s, u*.

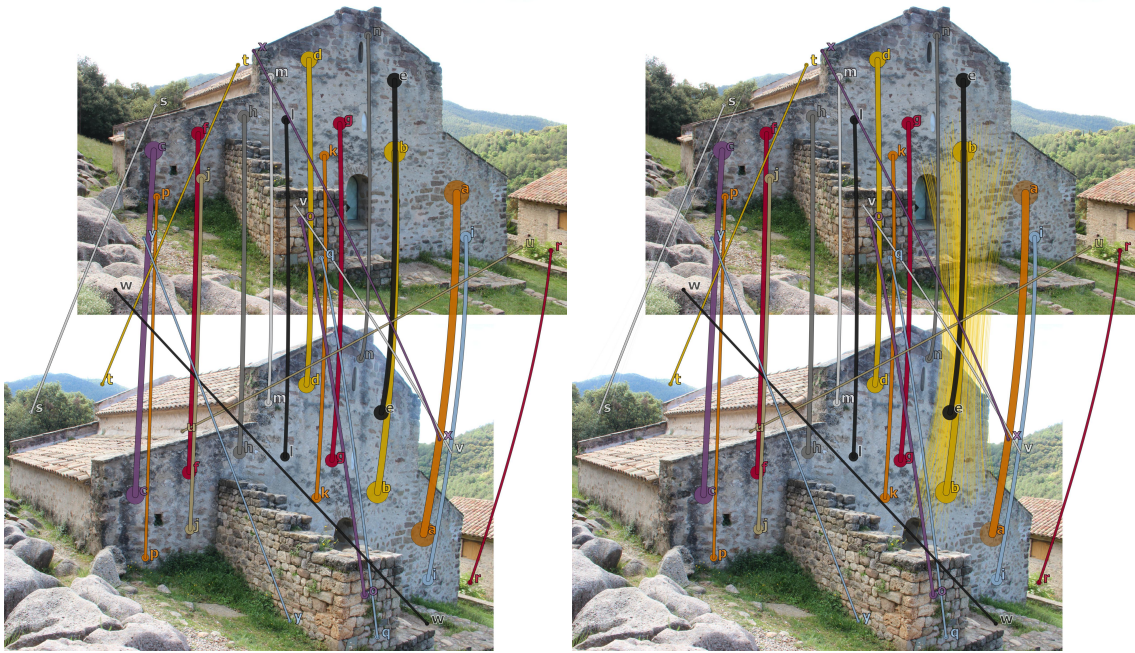
(b) Aggregated segments using *avg*. We can observe outliers that stand out such as *m, t, x*.



(c) Aggregated segments using *max*. We can observe outliers that stand out such as *y, x*.

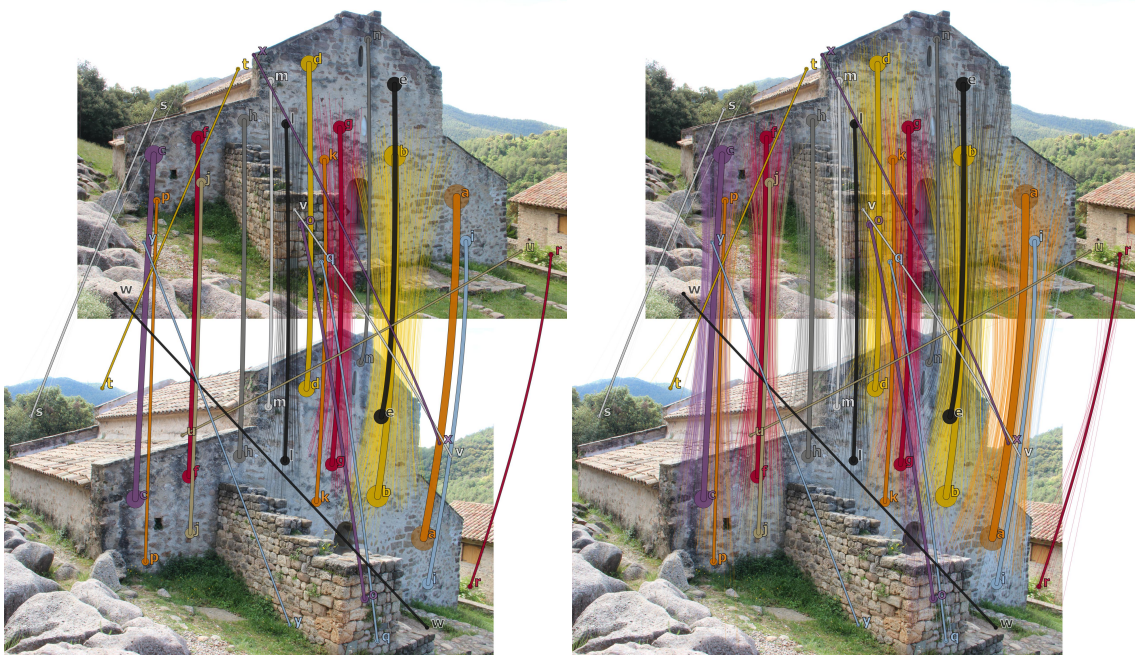
(d) Baseline approach. Just a few outliers can be distinguished.

Figure 2: Aggregated segments using *min*, *avg* and *max* methods on a facade, compared with a baseline showing all matches. Our output includes also glyphs to identify segments.



(a) Visualization generated with no uncollapsed segments.

(b) Visualization generated with a small amount of uncollapsed segments.



(c) Visualization generated with a moderate amount of uncollapsed segments.

(d) Visualization generated with a large amount of uncollapsed segments.

Figure 3: Benefits of hierarchical clustering: Church image pair with an increasing number of uncollapsed segments.



(a) Visualization generated with no uncollapsed segments.

(b) Visualization generated with a small amount of uncollapsed segment



(c) Visualization generated with a moderate amount of uncollapsed segment

(d) Visualization generated with a large amount of uncollapsed segment

Figure 4: Benefits of hierarchical clustering: image pair of Notre-Dame de l'Annonciation - Puy-en-Velay (courtesy of Zoilo Perrino), with an increasing number of uncollapsed segments.



Figure 5: Output with image pairs with large scale variations. Indeed, scale variations reduce the likelihood of getting roughly parallel segments. When combined with image layout optimization, resulting drawings have clearly distinguishable paths. Model: The dance of the Muses on Helicon, courtesy of Geoffrey Marchal.



Figure 6: Additional results with a facade and a decorated column. The facade case is specially challenging for the matching algorithm due to the amount of symmetries and repeated patterns they present. For the column case we again observe that the segments produced for images with different scales are not parallel. Column image credit: Owls (Louvre), Benoît Rogez.