FAsT-Match: Fast Affine Template Matching

Simon Korman, Daniel Reichman, Gilad Tsur, Shai Avidan

1 A. Proof of Theorem 3.1

We first restate Theorem 3.1 for completeness:

Let I_1, I_2 be images with dimensions n_1 and n_2 and let δ be a constant in $(0,1]$. For a transformation T' let T be the closest transformation to T' in a δn_1-cover $\mathcal{N}_\delta$. It holds that

$$|\Delta_{T'}(I_1,I_2) - \Delta_T(I_1,I_2)| \leq O\Big(\delta \cdot \frac{\mathcal{V}}{n_1}\Big)$$

To understand why the claim holds we refer the reader to Figure 1. Two close transformations T, T' map the template to two close parallelograms in the target image. Most of the error of the mapping T' is with respect to the area in the intersection of these parallelograms (the yellow region in Figure 1). This error cannot be greater than the total variation multiplied by the distance between the transformations T and T', as shown below. The rest of the error originates in the area mapped to by T' that is not in the intersection (the green region). The size of this area is also bounded by the distance between the transformations. Thus, the distance between the transformations, and the total variation, bound the difference in error between T and T'. This is formalized in the remainder of the section.

For convenience, throughout the discussion of the algorithm's guarantees we consider points in a continuous image plane instead of discrete pixels. Analyzing the problem in the continuous domain makes the theorem simpler to prove, avoiding several complications that arise due to the discrete sampling, most notable, that several pixels might be mapped to a single pixel. We refer the reader to a (slightly more involved) proof in the discrete domain, which we made available in a previous manuscript [1].

In order to switch to the continuous domain, we give some definitions and state some claims for points in the image plane. We begin by relating the intensity of points to that of pixels.

***Definition** 1.1:* The intensity of a point $p = (x, y)$ in the image plane (denoted $I_1(p)$) is defined as that of the pixel $q = ([x], [y])$, where $[\cdot]$ refers to the 'floor' operation. The point p is said to *land in* q.

We now define the variation of a point and relate it to the variation of a pixel.

***Definition** 1.2:* The variation of a point p, which we denote $v(p)$, is $\max_{q\,:\,d(p,q)\leq 1} |I_1(p) - I_1(q)|$. Note that this is upper-bounded by the variation of the pixel that p lands in. For convenience of computation (this does not change the asymptotic results), for points p that have a distance of less than 1 from the boundary of the image, we define $v(p) = 1$.

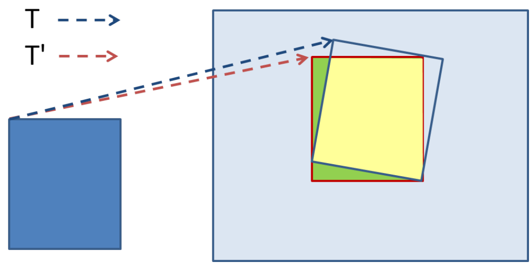

Fig. 1. **A template mapped to an image by two close transformations.** The close transformations map the template to close parallelograms. The error of T' cannot be very different from that of T. Most of the change in error is from different points being mapped to the intersection area (in yellow). This difference depends on the total variation of the template. The remaining error depends on the green area which is small because the transformations are close.

Finally, we define the total variation of an image in terms of the total variation of points in the image plane.

***Definition** 1.3:* The total variation of an image (or template) I_1 is $\int_{I_1} v(p)$. We denote this value $\mathcal{V}$. Note that this is upper bounded by the total variation computed over the pixels.

Our strategy towards proving Theorem 3.1 involves two ideas. First, instead of working with the pair of transformations T and T', we will more conveniently (and we show the equivalence) work with the identity transformation I and the concatenated transformation $T'^{-1}T$. Second, note that in Theorem 3.1, we bound the difference in error between transformations T and T', which are δn_1 apart. A simplifying approach, is to 'relate' the transformations T and T' through a series of transformations $\{T_i\}_{i=1}^m$ (where $T_0 = T$ and $T_m = T'$), which are each at most at a unit distance apart, with $m = O(\delta n_1)$. Thus, in Claim 1.3 we handle the case of transformations that are a unit distance apart.

In the following lemmas we introduce a constant u, such that if $\ell_\infty(T, T') \leq u$ it holds that $\ell_\infty(T^{-1}, T'^{-1}) \leq 1$.

***Claim** 1.1: Given affine transformations T, T' with scaling factors in the range $[1/c, c]$ such that $\ell_\infty(T, T') \leq \delta n_1$, it holds that $\ell_\infty(T^{-1}, T'^{-1}) = O(\delta n_1)$.*

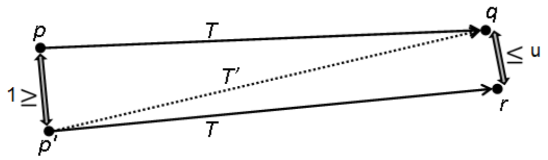

Fig. 2. **Illustration for Claim 1.1**. The distance between the points p and p' can be no more than a constant size greater than the distance between the points q and r, which is itself bounded by δn_1.

Proof: To see that the claim above holds, consider a point q and we will show that $||T'^{-1}(q) - T^{-1}(q)|| \leq c\delta n_1$ (see Figure 2). Let $p' = T'^{-1}(q)$ and let $p = T^{-1}(q)$. We wish to bound $||p' - p||$. Let $r = T(p')$. We get $||p' - p|| = ||T^{-1}r - T^{-1}q|| = ||T^{-1}(r - q)|| \leq c||r - q|| \leq c\delta n_1$. □

***Claim** 1.2: There exists a value $u \in (0,1)$ such that for any affine transformations T, T' where $\ell_\infty(T, T') \leq u$ and for any point $p \in I_1$, it holds that $||p, T'^{-1}(T(p))|| \leq 1$.*

The correctness of Claim 1.2 follows directly from Claim 1.1 by noting that $p = T^{-1}(T(p))$.

***Claim** 1.3: Let I_1, I_2 be images with dimensions n_1 and n_2. There exists a constant $u \in (0,1)$ for which the following holds. For any two affine transformations T and T' such that $\ell_\infty(T, T') \leq u$:*

$$|\Delta_{T'}(I_1, I_2) - \Delta_T(I_1, I_2)| \leq O\Big(\frac{\mathcal{V}}{{n_1}^2}\Big)$$

Note that the value $O\Big(\frac{\mathcal{V}}{{n_1}^2}\Big)$, bounds the difference in error for two transformations that have unit distance. This scales to the value $O\Big(\delta \cdot \frac{\mathcal{V}}{n_1}\Big)$ that appears in Theorem 3.1 for transformations that have a distance of δn_1.

Proof: Using the triangle inequality we can write:

$$\Big|\Delta_{T'}(I_1, I_2) - \Delta_T(I_1, I_2)\Big| = \Big|\int_{I_1} |I_1(p) - I_2(T'(p))| - \int_{I_1} |I_1(p) - I_2(T(p))|\Big| \leq \int_{I_1} ||I_1(p) - I_2(T'(p))| - |I_1(p) - I_2(T(p))||$$

where integrals go over points p in the template I_1.

We now bound this sum. As we know that $\ell_\infty(T, T') \leq u$, we know that only points that have a distance of at most 1 (as $u \leq 1$) from the boundary of I_1 are mapped to 'new' areas of I_2 - areas to which no point from I_1 was mapped before. Each of these points has an error of 1 at worst (this is the greatest distance possible between intensities from 0 to 1). The total area of such points is $O(n_1)$, and thus they contribute $O(n_1)$ to the difference between $\Delta_{T'}(I_1, I_2)$ and $\Delta_T(I_1, I_2)$, before normalization. This is equal to their contribution to the total variation.

For the remaining points (that have distance greater than 1 from the boundary of I_1), under T each such point p is mapped to a point $T(p)$, and the pre-image of that point $T'^{-1}(T(p))$, is in the area of I_1. Instead of considering the value $E_{T,I_1,I_2}(p)$ for each such point p in I_1, consider instead the error over each point $q = T(p)$ in I_2 that has points mapped to it both by T and by T'. The distance between p and $T'^{-1}(T(p))$ is at most 1 (as seen in Claim 1.2), and the value of p and of $T'^{-1}(T(p))$ differ by at most $v(p)$, and thus $|I_2(q) - I_1(p)| - |I_2(q) - I_1(T'^{-1}(q))| \leq v(p)$ (By a triangle inequality). Thus, for points that have a distance greater than 1 from the boundary of I_1, the affect on the difference $|\Delta_{T'}(I_1, I_2) - \Delta_T(I_1, I_2)|$ for each point p is at most $v(p)$ and thus the total contribution is bounded by $\mathcal{V}$.

Summing both contributions and normalizing by ${n_1}^2$ we get $|\Delta_{T'}(I_1, I_2) - \Delta_T(I_1, I_2)| = O(\mathcal{V}/{n_1}^2)$ as required. □

However, not all transformations have a distance of u from the net. We now turn to the goal of this section, proving Theorem 3.1.

Proof: As T is the closest transformation to T' it holds that $\ell_\infty(T, T') \leq \delta n_1$. Furthermore, from the construction that is summarized in Claim **??** we have that for $T = TrR_2SR_1$ and $T' \in \mathcal{N}_\delta$ where $T' = Tr'R_2'S'R_1'$ such that $d(Tr, Tr') \leq \delta n_1$, ... $d(R_1, R_1') \leq \delta n_1$. Now consider a series of transformations $\{T_i\}_{i=1}^m$ where $T_0 = T$ and $T_m = T'$. For each transformation T_{i+1} it will hold that $\ell_\infty(T_i, T_{i+1}) \leq u$ (for the constant u from Claim 1.3). For such a series, repeated use of Claim 1.3 (and of the triangle inequality) will give us that

$$|\Delta(T) - \Delta(T')| = |\Delta(T_0) - \Delta(T_m)| \leq O\Big(\frac{m\mathcal{V}}{{n_1}^2}\Big)$$

To construct such a series of transformations we first add (or subtract) u from the translation matrix until it changes from Tr to Tr'. This takes $O(\delta n_1)$ steps. We then change the rotation matrix beginning with R_2 by u/n_1 for $O(\delta n_1)$ steps until we get to R_2'. We proceed like this and after $m = O(\delta n_1)$ steps transition from T to T', giving us the required bound of $O\Big(\delta \cdot \frac{\mathcal{V}}{n_1}\Big)$. □

Acknowledgements This work was supported by the Israel Science Foundation (grant No. 873/08, in part) and the Ministry of Science and Technology.

References

[1] S. Korman, D. Reichman, and G. Tsur. Tight approximation of image matching. *CoRR*, abs/1111.1713, 2011.