HP - ICVGIP competition

Indian Institute of Technology
KHARAGPUR

Project title

"Automated Mosaicing of Torn Paper Documents"

By
Amiya Patanaik
Bibek Behera
Sukadeb Acharya

Project guide
Dr. Abhir Bhalerao
Department of Computer Science
University of Warwick – Coventry, UK

Introduction

Although the problem of torn paper reconstruction may sound same as the jigsaw problem, it is very different from it. The very process of tearing a paper causes shears and mechanical distortion which makes the problem much more difficult to handle compared to a jigsaw problem. Contrary to jigsaw problem here the assumption can be made that the torn pieces will be very different from each other in terms of shape. As a result only the matching of contours is sufficient to reconstruct the paper. Moreover the shears at the edges make it very difficult to come up with a simple continuity test at the corners being joined. Our initial research is based on contour scale space analysis (CSS) of the contour although the results are not yet satisfactory but we believe with continued experimentation we can perfect the method. For the competition our approach is similar to that of [2].

Scale Space Approach

We obtain the grayscale version of each image *I*. A Gaussian filter of standard deviation $\sigma = 3.0$ is applied to *I*. The result is then thresholded to obtain a mask *M* such as $M(x,y) = 0$ for background pixels and $M(x,y)=1$ for piece of paper. Morphological closing is done with a circular structuring element *B* of radius 2 pixels.

$$
\begin{aligned}
&L_g(x,y) = g(\sigma) * I(x,y), \\
&g(\sigma) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \\
&M(x,y) = \begin{cases} 0 \text{ for background pixels in } \mathrm{L}_g \\ 1 \text{ otherwise} \end{cases} \qquad (1.1) \\
&\hat{I} = M \circ B \text{ where } \circ \text{ is morphological closing operation}
\end{aligned}
$$

Now canny edge detection algorithm is applied to $\hat{I}$ and the result is skeletonised to obtain 1 pixel thick border. In case noise is present, small contours also come into picture which are removed by taking into consideration only the contour with maximal enclosing area. Curvature Scale Space is a technique for object representation, invariant under pose variations and based on the scale space.To build the CSS representation the curve needs to be considered as a parametric vector equation $\Gamma(t) = (x(t), y(t))$, then a series of evolved versions of $\Gamma(t)$ are produced by increasing the scale parameter, σ, from 0 to ∞. Every new evolved version is defined as $\Gamma_\sigma(t) = (X(\sigma,t), Y(\sigma,t))$, where

$$
\begin{aligned}
X(\sigma,t) &= x(t) * g(\sigma) \\
Y(\sigma,t) &= y(t) * g(\sigma)
\end{aligned}
\qquad (1.2)
$$

Since the CSS representation contains curvature zero-crossings or extrema points from the evolved version of the input curve, these are calculated directly from any Γ_σ by:

$$
\begin{aligned}
k(t) &= \frac{X'(t,\sigma)Y''(t,\sigma) - Y'(t,\sigma)X''(t,\sigma)}{(X'(t,\sigma)^2 + Y'(t,\sigma)^2)^{\frac{3}{2}}} \\
X'(t,\sigma) &= x(t) * \frac{\partial g(t,\sigma)}{\partial t} \\
X''(t,\sigma) &= x(t) * \frac{\partial^2 g(t,\sigma)}{\partial t^2}
\end{aligned}
\qquad (1.3)
$$

Similarly $Y'(t,\sigma)$ and $Y''(t,\sigma)$ can be computed.

Now $k(t)$ acts as the signature of the curve. It has the following properties.

1. Local
2. Rotationally and translation ally invariant
3. Stable in the sense those small changes in curve should have small effect.
4. Curve can be re-constructed from signature.

Now the $k(t)$ is sampled at regular intervals and quantized into appropriate pocket sizes. This list of $k(t)$ can now be considered as the DNA of the curve. Matching of contours can now be considered as matching of DNAs with penalty for gaps. So the problem is now converted to longest consecutive sequence problem with higher penalty for larger gaps. This can be solved using dynamic programming. The method described is minimalistic and over simplified. Results are quite satisfactory for the time we have spent on it. But further work must be done to get good results.

Turning Function Approach

The results of CSS approach are not yet good enough to be submitted to the competition. So we went for the turning function based approach. In fact turning function $\Theta(t)$ is closely related to $k(t)$

$$\Theta'(t) = k(t) \tag{1.4}$$

Turning function has the following properties

1. Satisfies the characteristics of a signature.
2. It starts from 0 for simplicity and reaches an angle of 2pi for closed curve.
3. Turning function calculates the difference between the angle the tangent makes with the contour at each point and adds it to the angle of the previous point to get the turning function at the subsequent point.

Matching, Transformation and Merging

Then this turning function is sampled at an interval δ and stored as a DNA pattern such as $\mathbf{U_1 = \varphi_1\varphi_2\varphi_3\ldots\ldots\varphi_n}$ .While comparing a pair of DNA pattern, the smaller one is reversed while the longer one is copied and placed at its tail, such that now the angle accumulated over the string is 4π.The shorter sequence is slided through the longer one and for every shift '*d*', the difference $\mathbf{\Delta\varphi_{ab}^{\ d} = \varphi_b^{\ d} - \varphi_a}$ is stored. Since the DNA signature increases monotonically, at the point where the pattern matches, the difference is nearly a constant. So we sample $\Delta\varphi_{ab}^{\ d}$ at constant intervals of $\mathbf{\Delta t=10^\circ}$ and we obtain the points in the DNA patterns which could be potential match. This is indicated by a variable called

Confidence = (length/no. of turns)* length (where length denotes length of matched string)

For a straight edge no. of turns is zero and is rejected. Among all combinations the one with highest confidence is declared a match. For this pair, we obtain a transform E in the following manner:-

Suppose the segments that matched are C_1, C2. We represent these curve segments by the actual points on the contour for the final image (rather than the polygon approximation).Let these strings be U_j, Vj (1<j<n).The problem is to find an Euclidean transform E that minimizes mean square error between EU_j and V_j (1<j<n).For simplicity, shift origin of U_j such that $\Sigma U_j = 0$.

Then the transformation (which has been calculated to give lowest mean square error) is applied

$\mathbf{E = R_\theta U + a}$

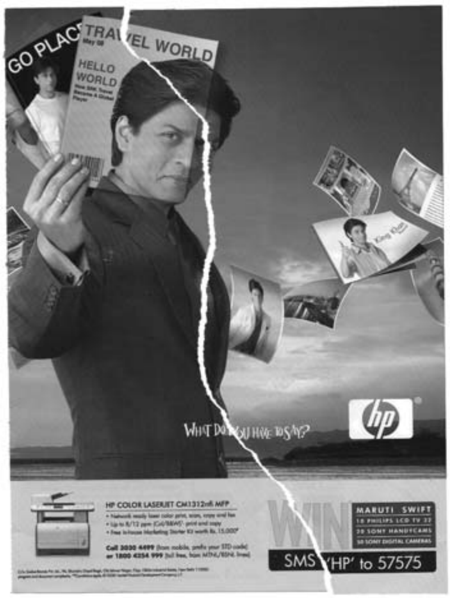
Where $a = (1/N)\, \Sigma V_j$ (for j between 1 and n) and R_θ = angle ($\Sigma U_j V_j$') or angle (ΣU_j Vj') + π such that intersection is lowest.

If there is no intersection between U and V then it is failure. We then select the second best pair of sequence as per confidence values. When we join the images the result is a single image. This process is run recursively until we obtain a single image which is the desired result.

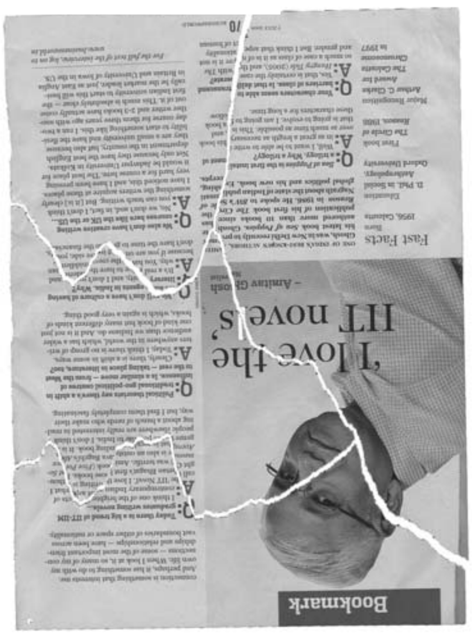
Problem of Global Consistency

We have completely neglected the problem of global consistency. Just because two pieces match doesn't mean they form the correct pair. Every merging operation must ensure that the joined peace form won't violate global consistency. One way to avoid the problem is to be able to backtrack to any merger step if that merger creates problem in future stages. This is computationally costly. As a temporary solution we have employed mode based operation. In *auto* mode all the pieces are given as input while in *batch* mode only two images are given at a time. So the user inputs two images which he feels have highest probability of merging. Then the merged output is fed along with another image. This process is continued till we end up with the complete reconstructed paper.

Sample 1



Sample 2



Sample 6

Sample 1 and Sample 2 reconstructed in *auto* mode. Sample 6 failed to reconstruct in *auto* mode but worked out in *batch* mode.

Future work

Although a good amount of time is devoted to the project. The approach is still minimalistic in nature and further work must be carried out to better it. One problem faced by this approach other than global consistency is that the system will fail to merge edges which are nearly straight. Also the lack of global consistency solution causes the output to be different for different order of input images. We believe that our CSS based approach can be a new way of handling the problem and with proper research can yield result as good as current methods if not better.

Sample 3

Poor result with pieces with nearly linear edges.

References

[1] Biswas, P. Bhowmick, and B. B. Bhattacharya, Reconstruction of Torn Documents Using Contour Maps, In Proc. ICIP 2005, III:517-520.

[2] L Zhu, Z Zhou and D Hu, Globally Consistent Reconstruction of Ripped-Up Documents, IEEE Transactions on Pattern Analysis and Machine Intelligence, Volume 30 , Issue 1 (January 2008)

[3] Wolfson, H.J, On curve matching, IEEE Transactions on Pattern Analysis and Machine Intelligence, Volume 12, Issue 5, May 1990 Page(s):483 – 489

[4] Prabhudev I. Hosur and Kai-Kuang Ma, OPTIMAL ALGORITHM FOR PROGRESSIVE POLYGON APPROXIMATION OF DISCRETE PLANAR CURVES

[5] Daniel Valdes and Abhir Bhalerao, Local Shape Modelling for Brain Morphometry using Curvature Scale Space