

A Fast and Stable Approach for Restoration of Warped Document Images

Kok Beng Chua, Li Zhang, Yu Zhang and Chew Lim Tan
School of Computing, National University of Singapore
3 Science Drive 2, Singapore 117543
{chuakokb, zhangli, zhangyu, tancl}@comp.nus.edu.sg

Abstract

We present a framework for acquiring and restoring images of warped documents to their original planar shapes. Image capturing of warped documents often results in warped images. Most digital restoration approaches make use of 2D image processing methods which are dependent on the content of the images. Some recent work attempts to remove this constraint by working directly on the 3D shape of the warped documents. In our framework, we will improve on this recent approach to make it more efficient and stable. First we capture the warped surface representation through a set of active lighting from a laser range scanner onto the document surface. This gives us an accurate 3D representation of the document shape. Next by using a stick constraint and a stable numerical integrator, we can digitally flatten the 3D model to a planar shape. This simple approach avoids potential instability problem and is fast. Restoration results for a few types of warped documents are demonstrated and a significant improvement over the original images can be observed.

1. Introduction

Printed materials are digitized for electronic dissemination and archival. Materials that can not be digitized using flat-bed scanners are often digitized by high quality digital cameras, producing 2D images. While imaging from a warped document, the curving of the document facing the camera causes a geometric distortion of the 2D content. While purely 2D image operations can improve the perceptual quality of the image, they can not address shape-based distortions without additional constraints. The pinhole camera is a projective transformation engine, and by its nature creates an image that is projectively distorted.

Zhang et al. proposed a resolution-dependent approach to straighten warped text lines in scanned document images on a word basis using polynomial

regression [1]. This approach requires the presence of single column text lines in the document. Zhang et al. also presented a method for image restoration by estimating the 3D surface shape of the document [2]. However, this is only applicable for warped documents which have a uniform cross-sectional area. Brown [3] proposed a novel approach of image restoration through the acquisition of the 3D geometry and restoring it using particle-based system and a mass spring model. However, the structure of the 3D shape is defined on a regular quad meshes which gives a less accurate representation of the true shape of the document. Furthermore, instability problems may arise when the wrong parameters are used in the mass spring model.

In this paper, we address three key issues in current image restoration approaches: flexibility, stability and speed. First, we acquire the 3D model of the warped document using a laser range scanner which gives us a very accurate representation of any 3D shape. Next, by using a downward force together with stick constraint, we can simulate the digital flattening of the document. The stick constraint eliminates the trouble of finding suitable spring constant in the mass spring model. The integrator used is a Verlet method which is more stable than the traditional simple Euler method. This allows us to use larger time-step thus speeding up the simulation.

The rest of this paper is organized as follows. Section 2 describes the acquisition of the 3D shape of the warped document. Section 3 illustrates the computer graphics techniques we used to flatten the non-planar model. Section 4 presents the restoration results of the various types of warped documents used in our experiment. Section 5 draws a conclusion and discusses some future works.

2. 3D Acquisition

The 3D models are acquired using the Minolta Vivid 900. The Vivid 900 uses LASER triangulation which returns very accurate 3D information. It is able to capture 307,000 points in only 2.5 seconds. Figure 1(a) shows the original mesh. To work on all these

points will need a lot of computational time even though it will return a very accurate result. We found in our experimental results that using just 1000 points is enough to give us an accurate simulation as shown in Figure 1(b).

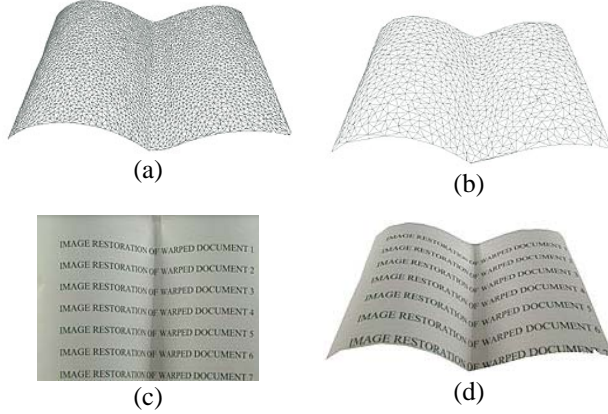


Figure 1. (a) Original 3D mesh (b) 3D mesh sub-sampled to 1000 points (c) 2D image (d) 3D mesh with 2D image as texture

Based on the triangulation information provided by the scanner, we can generate the 3D mesh for manipulation. The scanner returns images that are equivalent to a 3 CCD digital camera with full 24-bit color depth and an output pixel size of 640×480 . Figure 1(c) shows an example. These images are then registered onto the 3D model using the texture mapping information generated by the scanner as shown in figure 1(d).

3. 3D Model Restoration: Digital Flattening

Our framework makes use of a physically based modeling [4, 5] approach to perform the digital restoration of the warped document. A particle system is governed by the classic second order Newtonian equation, $f = ma$, where f is a force, m is the mass of a particle and a is its acceleration. The particle 3D space position can be represented by $[x_i, y_i, z_i]$ where i is the index of the particle. Since the document is considered as a rigid object, the 3D shape deformation can be defined by a distance constraint: any change in shape must preserve the distance between all points on the surface. Brown [3] proposed the use of a mass spring system [6, 7, 8, 9] to achieve the above objective. At each time-step, the springs attempt to restore its original length by using Hooke's law. However, choosing a suitable spring constant is a major problem [10]. Strong springs lead

to stiff systems of equations that result in instability if only simple integration techniques are used. On the other hand, weak springs make the 3D model too elastic. However, an interesting thing happens if we let the stiffness of the springs go to infinity. The system suddenly becomes solvable in a stable way with a very simple and fast approach since we can now exclude spring force computation.

3.1 Stick Constraint

Based on the above observation, we can model the mesh edge as a stick instead of a spring [10]. The stick can be thought of as infinitely stiff springs between each pair of particles. These springs are so strong that they will restore instantly to its original rest length from its deformed state. We simulate the stick by constraining the two particles to have a fixed distance between them.

This constraint can be expressed mathematically by the following equation:

$$(p[1] - p[0]) \cdot (p[1] - p[0]) = r^2 \quad (1)$$

where $p[1]$ and $p[0]$ are the current positions of two particles connected by the stick, r is the rest length of the stick.

During the simulation, the particles are pushed or pulled away to obtain the correct distance. Refer to Figure 2 for illustration.

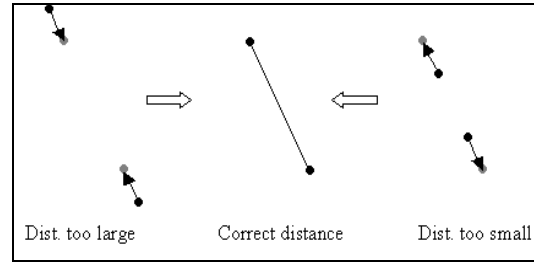


Figure 2. Fixing an invalid distance

3.2 External Forces

3.2.1 Global Downward Force. Our main goal in the simulation is to drive all the particles down to a plane. To achieve this, we need a global downward force that is exerted on all particles to force them toward the plane. This force is typically modeled as: a constant force, f , gravity and can be represented as:

$$f = mg \quad (2)$$

where m is the mass of the particle and g is the gravitational acceleration.

3.2.2 Plane Collision. One of the goals of the restoration process is to “flatten” the 3D model by driving all the particles to the $Z=0$ plane. This plane defines the plane which the warped document was placed onto during the 3D shape acquisition phase. During the flattening process, some particles may collide with or penetrate through this plane.

In Brown’s implementation [3], collisions are handled by changing the velocity and direction of the particle. This can be done by adding a coefficient of restitution. If a particle penetrates through the $Z=0$ plane, the system is restored to the previous time and a new time-step is calculated. This is not very practical from a real-time point of view since the simulation could potentially run very slowly when there are a lot of collisions or penetrations.

Here, we use yet another strategy. Offending points are simply projected out of the obstacle. By projection, we mean moving the point as little as possible until it is free of the obstacle. This means moving the point perpendicularly out towards the $Z=0$ surface. This can be defined as follows:

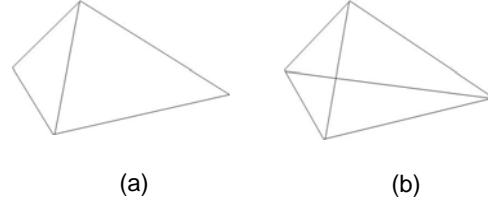
$$[x_i, y_i, z_i] = [x_i, y_i, \max(z_i, 0)] \quad (3)$$

By satisfying the above plane constraint, we may end up invalidating the stick constraint. We will look at how we can remedy such a situation in the later section.

3.3 Mesh Reconstruction

Unlike the regular mesh [3, 11] that Brown uses for his simulation, the 3D scanner can only return an irregular triangular mesh as shown in Figure 1.

Regular quad mesh may increase the stability of a simulation but is a big constraint when modeling complex shape. It can not accurately represent complex shapes. In such situations, an irregular mesh is more suitable. However, in our experiment, we found that it gives rise to instability problem. The mesh is unable to hold the shape of the 3D model after the flattening process. The irregular mesh configuration makes it less rigid. We solve this problem by adding a bending resistance [6, 12, 13] to increase the rigidity of the mesh. A bending resistance is added for every pair of triangles that share an edge as shown in Figure 3. These extra sticks help to maintain the shape of the 3D model by preserving the space between diagonal elements.



**Figure 3. (a) Initial triangles sharing one edge
(b) bending resistance added**

3.4 Numerical Simulation

In most particle based systems, we can see that each particle has two main variables, the position, x and its velocity, v .

Brown [3] proposes the use of simple Euler method to handle this system. Then in each time-step, the new position x' and v' are calculated as follows:

$$v' = v + a\Delta t \quad (4)$$

$$x' = x + v'\Delta t \quad (5)$$

where Δt is the time step, and a is the acceleration computed using classic second Newtonian equation, $f = ma$. In the Euler algorithm, mostly due to floating point errors, the x and v can get out of sync because they are stored separately. This can cause numerical instability and, in cases, a breakdown of the system. Thus to avoid this problem, we must take very small time-step which results in an increase of computational cost.

In order to increase the numerical stability, we use a different representation and integration scheme that Brown used. Instead of storing each particle’s position and velocity, we store its current position x and its previous position x^p . Keeping the time step fixed, the update rule (or integration step) is then:

$$x' = 2x - x^p + a \cdot \Delta t^2 \quad (6)$$

$$x^p = x \quad (7)$$

This is called Verlet integration [10, 14] and is used extensively when simulating molecular dynamics. It is quite stable since the velocity is implicitly given and consequently it is harder for velocity and position to come out of sync. It works due to the fact that $2x - x^p = x + (x - x^p)$ and $x - x^p$ is an approximation of the current velocity.

The projection of the particle back to the surface may invalidate the stick constraint. By satisfying the stick constraint it may cause the particle to penetrate through the plane again. This problem can be remedied by solving both constraints at the same

time. This is simply a case of solving a system of equations. However, we choose to proceed indirectly by local iteration. By solving both constraints for a number of times, we can see that it will slowly converge to a reasonable good result.

The number of necessary iterations varies depending on the physical system simulated and the amount of motion. While this approach of pure repetition might appear somewhat naïve, it turns out that it actually converges to the solution that we are looking for. It works by consecutively satisfying various local constraints and then repeating; if the conditions are right, this will converge to a global configuration that satisfies all constraints at the same time. By stopping the iterations early, one can trade off speed for accuracy.

The 3D document model is initialized with the geometrical data obtained from the 3D scanner followed by a restructuring of the model to increase stability. Each particle mass is set to 1 and a downward force of gravitational acceleration of 9.81, is then applied to “flatten” the model. The process is complete when all the particles are on the Z=0 plane.

4. Results

We tested this framework with two different kinds of distortion. For all the experiments, the 3D models are acquired using the Minolta Vivid900 (as described in section 2). We also apply a document binarization [15] process to each restored images to lessen the shading effects. This allows us to see more clearly the improvement after the restoration process.

4.1 Experiment I: Thick bound book

The first experiment deals with a distortion that can be commonly found in a thick bound book with a spine in the middle. Image capturing of such document type often results in a warped image of the book surface. Figure 4(a) shows a 2D image of the warped document. Figure 4(b) is the corresponding binary image of the warped document. Figure 4(c) shows the result after applying our restoration routines. Figure 4(d) is the corresponding binary image of the restored image. We can clearly observe that the curved text lines are now more straightened.

4.2 Experiment II: A warped document

In the second experiment, we attempt to restore a less commonly found distortion. A piece of paper folded by hand to obtain multiple folds as shown in Figure 5(a). We then try to remove the distortion caused by the multiple folds using our restoration

framework. Figure 5(b) shows a crop of the distorted region. We can clearly see that the curvature of the text caused by the fold-lines.

Figure 5(c) shows the results after applying our image restoration routines. We can observe that the distorted lines and texts are more straightened. However, the shading effects still remain. This makes the image look slightly more warped than it really is. Figure 5(d) shows the corresponding binary image of the restored image.

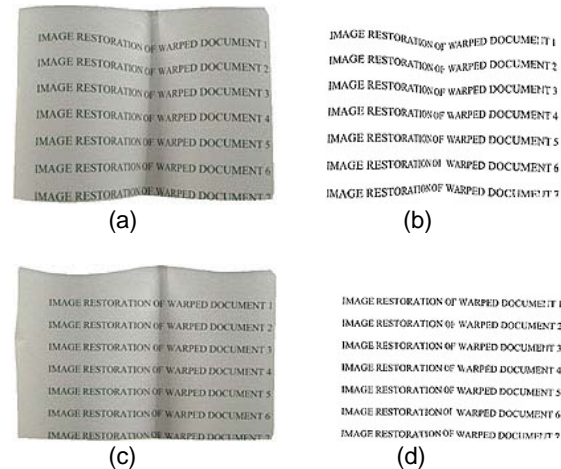


Figure 4. (a) Warped document (Grayscale) (b) Warped document (Binary) (c) Restored document (Grayscale) (d) Restored image (Binary)



Figure 5. (a) Folded paper (b) Portion of the distorted image (c) Restored result (d) Restored result (Binary)

4.3 Simulation Time

Both experiments are run on a Dell Intel Pentium III 996Mhz (512 MB RAM). There are 1000 points in each mesh. Time-step used is 0.1. Table 1 shows the summary of the simulation time of both experiments.

Table 1. Summary of simulation time

Experiment No.	Time to add bending resistance (seconds)	Time to flatten 3D model (seconds)	Total Time (seconds)
I	10	18	28
II	10	20	30

5. Conclusion and Future Work

In this paper, we present a framework for the restoration of warped document image using physically-based modeling. Though this idea is not new, we have made three key improvements in the following aspects: flexibility, stability and speed. First, since our framework is able to handle irregular triangular mesh, we can remove the constraint imposed on the complexity of the shape. Second, by using a stick constraint model, we are able to eliminate potential instability problems due to choosing incorrect parameters for the mass spring model. Last but not least, by using Verlet integration, we are able to use larger time-step which speeds up the numerical simulation.

In our paper, we do not handle shading removal. Shading is a strong visual cue for shape. Correcting geometric distortion without addressing shading artifacts may make restored images still appear distorted. Future work may include ways to remove the shading by analyzing the geometrical structure of the 3D model. To improve numerical stability further, we may also consider adopting an implicit numerical integration method. Our current implementation does not handle self-collision of the particles. This may pose a problem for warped documents that contain self occlusions. To solve this problem, we can use collision detection method to prevent self-intersection of the 3D mesh.

Acknowledgement

This research is supported in part by National University of Singapore URC grant R252-000-202-112 and Agency for Science, Technology and Research (A*STAR) grant R252-000-206-305.

References

- [1] Z. Zhang, Tan, C. L. Correcting document image warping based on regression of curved text lines. International Conference on Document Analysis and Recognition, Vol. 1, pp. 589-593, August 2003.
- [2] Z. Zhang, C.L. Tan, L. Fan. "Restoration of Curved Document Images through 3D Shape Modeling". *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'04)* Volume 1, pp. 10-15, July 2004.
- [3] M.S. Brown and W.B. Seales. "Document restoration Using 3D shape: A General Deskewing Algorithm for Arbitrary Warped Documents". *International Conference on Computer Vision (ICCV'01)* Volume 2, pp. 367-374, July 2001.
- [4] A. Witkin and D. Baraff. "Physically Based Modeling: Particle System Dynamics". In *SIGGRAPH*, 1997.
- [5] A. Witkin and D. Baraff. "Physically Based Modeling: Rigid Body Simulation". In *SIGGRAPH*, 1997.
- [6] G. Oliveira. "Exploring Spring Models". Game Developer. October 5, 2001.
- [7] J. Lander. "Devil in the Blue-Faceted Dress: Real-Time Cloth Animation". Game Developer. March 27, 2000
- [8] X. Provot. "Deformation constraints in a mass spring model to describe rigid cloth behaviour". In *Graphics interface*, pp. 155-174, 1995.
- [9] W. B. Seales and Yun Lin. "Digital Restoration using Volumetric Scanning". *ACM/IEEE-CS Joint Conference on Digital Libraries*, pp. 117-124, 2004.
- [10] T. Jakobsen. "Advanced Character Physics". Game Developer. January 21, 2003.
- [11] D. Baraff and A. Witkin. "Large Steps in Cloth Simulation". In *Computer Graphics (Proc. SIGGRAPH)*, pp. 43-52. August 1988.
- [12] K.J. Choi and H.S. Ko. "Stable but responsive cloth". In Conference Proceedings of SIGGRAPH 2002, pp. 604-611, 2002.
- [13] K. Choi and H. Ko. "Extending the Immediate Buckling Model to Triangular Meshes for Simulating Complex Clothes". In *EUROGRAPHICS*, pp. 187-191, 2003.
- [14] J. Dummer. "A Simple Time-Corrected Verlet Integration Method". Game Developer. 2004.
- [15] Z. Zhang, C.L. Tan, "Restoration of images scanned from thick bound documents", *International Conference on Image Processing*, 7-10 October 2001.