An Improved Physically-Based Method for Geometric Restoration of Distorted Document Images

Li Zhang, Yu Zhang, *Member*, *IEEE*, and Chew Lim Tan, *Senior Member*, *IEEE*

Abstract—In document digitization through camera-based systems, simple imaging setups often produce geometric distortions in the resultant 2D images because of the nonplanar geometric shapes of certain documents such as thick bound books, rolled, folded, or crumpled materials, etc. Previous work [1], [2], [3], [4] has demonstrated that arbitrary warped documents can be successfully restored by flattening a 3D scan of the document. These approaches use physically-based or relaxation-based techniques in their flattening process. While this has been demonstrated to be effective in rectifying the image content and improving OCR, these previous approaches have several limitations in terms of speed and stability. In this paper, we propose a distance-based penalty metric to replace the mass-spring model and introduce additional bending resistance and drag forces to improve the efficiency of the existing approaches. The use of Verlet integration and special plane collision handling schemes also help to achieve better stability without sacrificing efficiency. Experiments on various document images captured from books, brochures, and historical documents with arbitrary warpings have demonstrated large improvements over the existing approaches in terms of stability and efficiency.

Index Terms—Warped image restoration, geometric correction, physically-based modeling, numerical integration.

1 Introduction

Document digitization has been an important task to integrate various physical documents into the vast online database that is searchable, browsable, and readable by a large population worldwide. Current high-resolution digital cameras have made digitization tasks much simpler and more document-friendly with their noncontact imaging property. Due to the nonplanar geometric shapes of some documents and the pinhole camera's perspective projection property, many camera images often display geometric and perspective distortions [3], [5]. This creates problems for both human perception and subsequent document image analysis tasks. The objective of dewarping approaches is therefore to remove these distortions and produce a planar view of the warped document to preserve image fidelity and facilitate information extraction and further content analysis tasks.

Many recent approaches have tried to use the surface shape of the document to model the geometric warping, which is then rectified using physically-based or relaxation-based techniques through a flattening process. Such 3D-based approaches often produce better results than 2D-based approaches [6], [7], [8] due to more accurate representations of the physical warpings and are widely applicable to images with diverse content. More complex distortions such as folds and rumples are also manageable using 3D-based approaches as long as an accurate surface shape can be obtained. To get the surface shape, one direct way is through special setups such as structured lighting [3], [9], laser range scans [10], Eye Scanner [11], or Computed Tomography (CT) scans for opaque objects [12], [13]. For example, Pilu [1] presents a novel method that models curled paper deformations using an applicable surface and flattens it to a plane during restoration. Brown and Seales [2], [3] propose an approach that captures the 3D geometry of the document using a structured light acquisition system and then flattens it numerically through a particle-based mass-spring model. This method effectively restores arbitrary geometric distortions and is successfully applied to the digitization of damaged manuscripts in digital libraries. Sun et al. [9] also apply a similar geometric correction method to restore multi-folded art materials with a novel photometric correction method based on the notion of intrinsic images. Besides flattening the surface using iterative-relaxation techniques, Brown and Pisula [4] also propose to use conformal mapping to parametrize a document's 3D shape to a 2D plane by solving a large linear system, which eliminates the instability problems in previous relaxation-based methods. Furthermore, Seales and Lin present a method to reveal invisible texts buried in damaged books or scrolls [12], [13]. Unlike the previous approaches which mainly focus on opened documents, this method considers impenetrable objects by capturing the volumetric data through CT scans. Different from the 2D-based approaches, most of these 3D-based approaches do not depend on the content of the document though they might require a special setup to capture the 3D geometry. On the other hand, to avoid the use of special setups, various methods have attempted to employ shape recovery techniques to reconstruct the surface shape of the document, such as methods using Shape-from-Stereo [14] and Structure-from-Motion [15].

In this paper, we present an improved physically-based method for restoring geometrically distorted document images with greater efficiency and stability. The improvements enhance usability of 3D geometry-based document restoration and make such a system more practical. An overview of the work flow is as follows: First, a 2D image is captured using a laser range scanner which also returns the 3D range data representing the surface shape of the document. Next, the warped surface is modeled as a particle system and flattened through a numerical simulation process. By using the original image as the texture to the 3D surface, we obtain the geometrically restored image after flattening. The main contributions of our system are:

1. The geometric correction method uses a distance-based penalty metric, introduced as the stick constraint, to replace the mass-spring model used in many earlier approaches. This avoids the manual adjustment of the spring stiffness and damping coefficients which may cause problems when not properly tuned.
2. The use of drag forces in the geometric correction method helps to spread out the mesh during flattening, which avoids particle self-collisions and also speeds up the simulation process significantly.
3. Stability and efficiency are achieved in the geometric correction method by using Verlet integration with special plane collision handling schemes.

2 2D Image and 3D Shape Acquisition

We use a dedicated laser range scanner—Minolta Vivid 900 [16] to capture the 3D structure of the warped document. The object is scanned by a plane of laser light coming from the Vivid's source aperture. The scanner captures range scan and color data simultaneously and has a sample pitch of 0.068 mm within an optimal measuring range from 0.6 m to 1.2 m. Therefore, it produces an accurate triangular mesh together with a texture image of size 640×480 as shown in Figs. 1a and 1b, respectively. The laser range scanner is able to capture 307,000 points with an output image of size 640×480 in 2.5 seconds. However, working with all the 307,000 points is rather computationally expensive. Our experiments show that using

- *L. Zhang and C.L. Tan are with the School of Computing, National University of Singapore, 3 Science Drive 2, Singapore 117543. E-mail: {zhangli, tancl}@comp.nus.edu.sg.*
- *Y. Zhang is with the Institute of High Performance Computing, A*STAR, #01-01, 1 Science Park Road, The Capricorn, Science Park II, Singapore 117528. E-mail: zhangy@ihpc.a-star.edu.sg.*

Manuscript received 20 July 2007; revised 10 Oct. 2007; accepted 26 Oct. 2007; published online 21 Nov. 2007.
Recommended for acceptance by E. Saund.
For information on obtaining reprints of this article, please send e-mail to: tpami@computer.org, and reference IEEECS Log Number TPAMI-2007-07-0440.
Digital Object Identifier no. 10.1109/TPAMI.2007.70831.

0162-8828/08/$25.00 © 2008 IEEE Published by the IEEE Computer Society

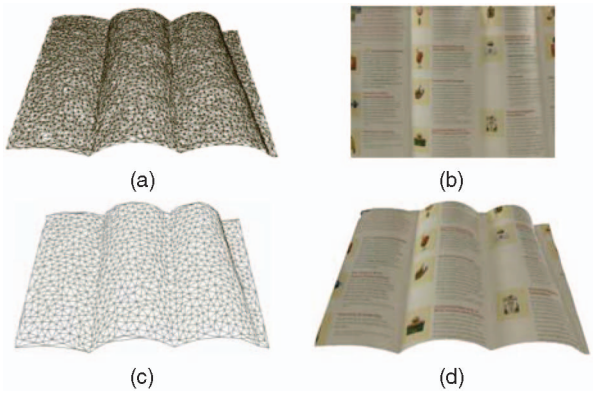

Fig. 1. Acquired document data: (a) Original 3D mesh. (b) Two-dimensional texture image. (c) Subsampled sparse mesh with 1,010 points. (d) Textured 3D mesh.

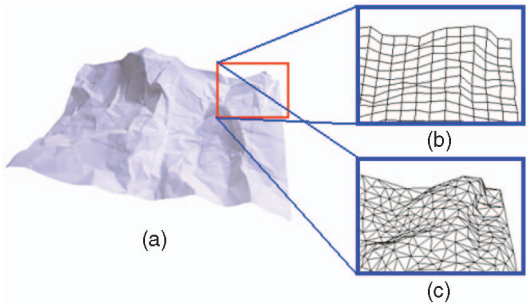

Fig. 2. The shape of a crumpled paper represented using an irregular triangular mesh and a uniformly sampled quad mesh (both have 1,200 points): (a) Original shape with texture. (b) Uniformly sampled quad mesh (partially enlarged). (c) Irregular triangular mesh (partially enlarged).

about 1,000 points is enough to produce a satisfactory restoration result. The subsampling process is done using the Minolta Polygon Editing Tool, which takes approximately 30 seconds to complete on a Pentium III 996 MHz PC. Figs. 1c and 1d show the subsampled mesh and the textured appearance, respectively.

3 Restoration Using Physically-Based Modeling

The restoration process is done using a physically-based modeling technique [17], [18]. These physically-based models use the classical Newtonian physics to govern the behavior of the 3D structural models represented as particle systems. They are often used in computer animation and computer graphics applications to simulate object deformations by adding controls on certain physical quantities such as forces, velocities, accelerations to the geometry, and tracking the changes of the geometry in response to the change of the physical properties. As discussed by Brown and Seales [3], the geometric correction problem can be considered as the reverse problem of sheet deformations. Starting with a deformed sheet which models the warped document surface, we need to reverse the deformation process and derive its planar state. The goal is to get a restored image with a better perceptual quality and also to facilitate further document analysis tasks.

3.1 Particle System Modeling

As discussed in Brown and Seales' paper [3], given the 3D geometry of the surface, a particle system is constructed with each point on the mesh as a particle. The whole system is governed by the classic second order Newtonian equation, $f = ma$, where f is the force exerted on the particle, m is the mass of the particle $(m = 1)$, and a is the acceleration. Each particle is associated with two variables, position x and velocity v. During the simulation process, dynamic forces are exerted onto the particles over time and make them move to new positions. Meanwhile, damped virtual springs connect

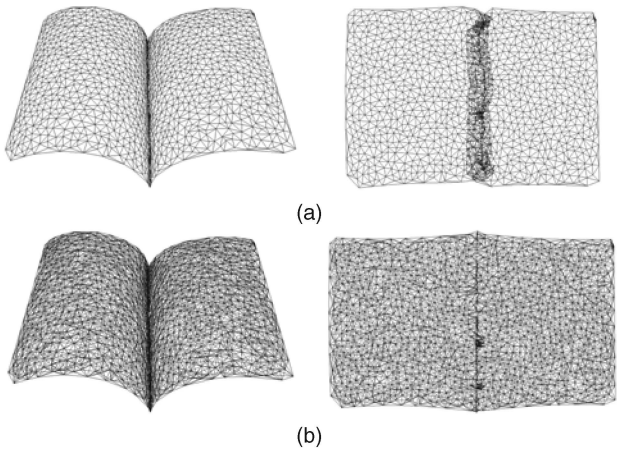

Fig. 3. (a) Flattening without bending resistance results in self-folding. (b) Flattening with bending resistance spreads out the mesh nicely especially near the spine region.

particles to impose distance constraints [19], [20]. The magnitude of the spring force is computed using Hooke's Law. At each time step, the spring forces are induced when external forces begin to move each individual particle from its rest state. These forces act together to keep the particles at a correct distance from each other and give the mesh its rigidity.

However, it is often difficult to choose suitable spring stiffness and damping constants. Strong springs lead to stiff meshes, while weak springs make the meshes too elastic. In our approach, instead of using springs, we use virtual sticks to connect the particles. These sticks can be considered as springs with infinite stiffness. In this case, we can exclude the spring force computation in favor of a point relaxation solver scheme [21].

3.2 Mesh Refinement

In our method, we use an irregular triangular mesh to represent the 3D geometry of the document. In general, irregular triangular meshes are more desirable than uniformly sampled quad meshes because more points can be allocated at places with more severe distortions. Therefore, irregular triangular meshes usually provide more accurate representations for the various surfaces we are dealing with. This was also brought out by Choi and Ko [22] in their simulation of complex cloth. Fig. 2 shows a comparison of using an irregular triangular mesh and a uniformly sampled quad mesh to represent the shape of a crumpled paper. Both meshes have 1,200 points, which is much less than the original dense mesh captured by the range scanner. The marked partial image shows that the irregular triangular mesh is more delicate in representing some fine rumples. Moreover, the subsampled sparse triangular mesh with 1,200 points does not degrade much from the original dense mesh. The measured distance between the sparse mesh and the dense mesh in terms of the depth Z for the crumpled page image is only about 0.38 mm on average and the standard deviation is 0.623 mm with a maximum depth of 64.84 mm.

We found in our experiments that an irregular triangular mesh may cause instability problems during flattening. The mesh may bend along the common edge of two triangles and result in a self-folding effect near some regions with large curvatures. To solve this problem, we propose to add a bending resistance between any pair of triangles that share a common edge. This increases the rigidity of the mesh and avoids the self-folding phenomenon as shown in Fig. 3.

3.3 Constraints and External Forces

3.3.1 *Stick Constraint*

As mentioned earlier, instead of using springs to connect the particles, we use sticks with fixed lengths. This is to avoid the spring force computation which is often error-prone due to the difficulty of

choosing suitable spring stiffness and damping constants. The sticks between the particles are simulated using the constraint that enforces a fixed distance between any pair of particles connected by an edge. This distance is defined as the initial euclidean distance between any two points on the original mesh. During the simulation, the particles are either pushed or pulled away in order to maintain the correct distance that is the stick's rest length. Mathematically, suppose A and B are the two end points of a stick with their positions at the time step n denoted as $P_a^n = (x_a^n, y_a^n, z_a^n)$ and $P_b^n = (x_b^n, y_b^n, z_b^n)$. The stick's rest length is defined as

$$d_{ab}^0 = \sqrt{(x_a^0 - x_b^0)^2 + (y_a^0 - y_b^0)^2 + (z_a^0 - z_b^0)^2}, \quad (1)$$

where $P_a^0 = (x_a^0, y_a^0, z_a^0)$ and $P_b^0 = (x_b^0, y_b^0, z_b^0)$ are the initial positions of A and B. Due to the external forces, A and B will move to positions $P_a^{n'}$ and $P_b^{n'}$ defined as

$$\begin{aligned} P_a^{n'} &= 2P_a^n - P_a^{n-1} + a \cdot \Delta t^2, \\ P_b^{n'} &= 2P_b^n - P_b^{n-1} + a \cdot \Delta t^2. \end{aligned} \quad (2)$$

To maintain the stick's rest length d_{ab}^0, the two end points A and B are thus moved away or toward each other by the same distance to reach their new positions P_a^{n+1} and P_b^{n+1} at the time step $n+1$

$$\begin{aligned} P_a^{n+1} &= P_a^{n'} - \frac{0.5(P_a^{n'} - P_b^{n'})(d_{ab}^n - d_{ab}^0)}{d_{ab}^n}, \\ P_b^{n+1} &= P_b^{n'} + \frac{0.5(P_a^{n'} - P_b^{n'})(d_{ab}^n - d_{ab}^0)}{d_{ab}^n}, \end{aligned} \quad (3)$$

where $d_{ab}^n = \sqrt{(x_a^n - x_b^n)^2 + (y_a^n - y_b^n)^2 + (z_a^n - z_b^n)^2}$.

3.3.2 Plane Collision Constraint

The goal of the flattening process is to drive all the particles on the 3D mesh to the $z = 0$ plane. This plane defines where the warped document is placed at the shape acquisition phase. During flattening, some particles might collide with this plane or penetrate through it. In Brown and Seales' work [3], this situation is handled by changing the velocity of the particle at the time of collision into the normal component and the tangential component. The normal component is then dampened with coefficient of restitution. However, this may cause the simulation to run very slowly when there are lots of collisions and penetrations. In our method, we introduce a dynamic inverse constraint on the displacement of the particle [20]. Particles that penetrated through the $z = 0$ plane will be projected out [21], i.e., move the particles perpendicularly toward the $z = 0$ plane as defined by: $(x_i, y_i, z_i) = (x_i, y_i, max(z_i, 0))$, where (x_i, y_i, z_i) is the position of the particle at the time of collision. This method of resolving collision allows a larger time step to be taken during the integration phase and, hence, speeds up the simulation significantly.

However, by satisfying the plane collision constraint, we might end up invalidating the stick constraint and vice versa. This can be resolved by solving both constraints simultaneously using a local iteration scheme known as point relaxation [21]. By repeatedly satisfying various local constraints, the whole system will eventually converge to a global configuration that satisfies all the constraints. The system terminates when the change of the surface area between time t and $t + \Delta t$ is less than a predefined ϵ so that further significant improvement is unlikely. In general, the number of iterations usually depends on the particle system being simulated and the amount of motion needed to reach the planar state.

3.3.3 Gravitational Force

To drive all the particles down to a plane, we have proposed to exert a downward gravitational force on all the particles to force them to move toward the $z = 0$ plane as in [10]. This force is a constant since $f = ma$, where $m = 1$ and $a = g = 9.81$.

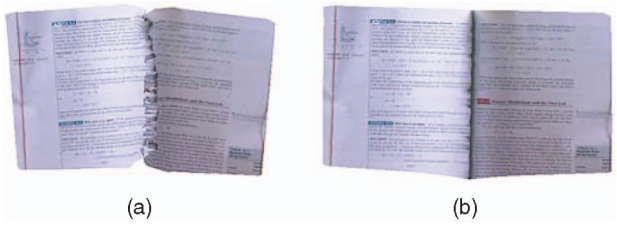

Fig. 4. (a) Restored image with gravitational force alone. (b) Restored image with additional drag force.

3.3.4 Drag Force

Sometimes using gravitational force alone is unable to simulate the deformation process correctly. This is especially true when the document warping is too severe like the deep spine in some thick bound books. In such cases, the mesh points near the spine region often collide with each other during the flattening phase even with an added bending resistance. This results in a self-collapsed spine as shown in Fig. 4a. Therefore, actions need to be taken to avoid such self-collisions and interpenetrations. Many studies on self-collision handling have been targeted to cloth simulation and garment design [23]. However, despite the many optimization techniques such as those based on the curvature of the surface [24] and bounding volume hierarchies [25], self-collision handling is still hard to achieve with real-time performance. Nevertheless, our objective here is mainly to obtain a final flattened state of the 3D model without caring about the exact deformation undergone at each time frame. Therefore, instead of detecting the self-collisions, we try to prevent them from happening. Volino and Magnenat-Thalmann [23] have pointed out that self-collisions only can happen at areas with large curvatures such as the steep spines or folds as in our case. Therefore, to avoid self-collisions at these regions, we add an additional drag force on each particle in both the horizontal and the vertical directions to force the particles to spread out during the simulation process. This drag force is similar to the viscous drag used in many particle systems [26] except that it is not to resist the motion but to slightly accelerate the motion as defined by

$$\begin{aligned} f_{drag}^h &= k_{drag} \cdot x_i^n, \\ f_{drag}^v &= k_{drag} \cdot y_i^n, \end{aligned} \quad (4)$$

where k_{drag} is a constant known as the coefficient of drag and (x_i^n, y_i^n) is the current 2D location of the particle i at the time step n. An empirical value of k_{drag} is chosen to be 0.05 in all our experiments and does not require special tuning. The origin of the particle system is defined as the center of the 3D shape and is therefore consistent across different inputs. Fig. 4b shows the geometrically corrected image with both the bending resistance and the drag force added.

3.4 Numerical Simulation

As mentioned earlier, each particle is associated with two variables, position x and velocity v. In Brown and Seales' approach, they use the simple explicit Euler method to handle the particle system [2], [3]. However, as discussed by Witkin [26], Euler's method can be unstable, inaccurate and inefficient. Due to errors in the formulation, x and v often become asynchronous as the time accumulates. This causes inaccuracy and instability problems and need to be avoided by taking small time steps. Though Euler's method requires only one evaluation per step, the small step size is a bottleneck that restrains its overall performance. Therefore, the use of Euler's method results in a low-computational efficiency. On the other hand, the error introduced by Euler integration is given as $-\frac{1}{2}a\Delta t^2 + O(\Delta t^3)$, which is much larger than other integration schemes such as Verlet integration and Runge-Kutta integration.

To increase numerical stability and to speed up the simulation, we use the Verlet integration [27] scheme here. Instead of storing both position and velocity, we only store the particle's current

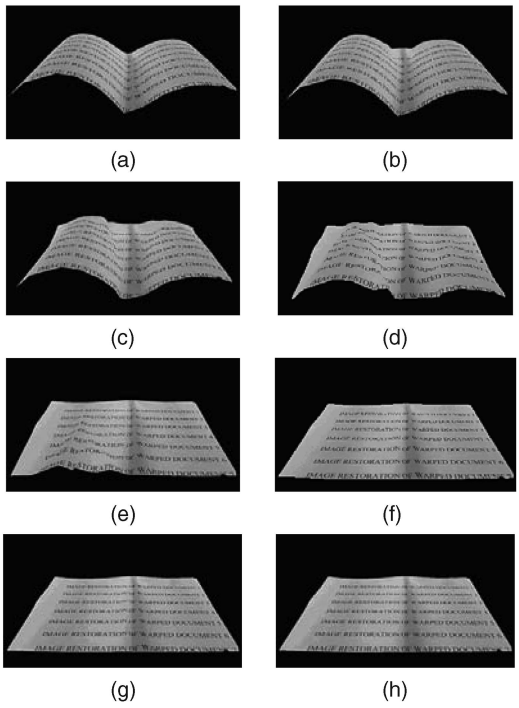

Fig. 5. Simulation of the flattening process showing eight frames of the textured mesh.

position x_n and its previous position x_{n-1}. The new position x_{n+1} can be calculated as

$$x_{n+1} = 2x_n - x_{n-1} + a \cdot \Delta t^2. \quad (5)$$

Verlet integration has been used extensively to simulate molecular dynamics. One example is to create water ripple effects in computer graphics simulation. This scheme is stable because the velocity is implicitly given as

$$v_n = \frac{x_{n+1} - x_{n-1}}{2\Delta t}. \quad (6)$$

Now, only the position of the particle needs to be taken care of. A large time step can thus be used to speed up the simulation process. In addition, compared to Euler integration, the error produced by Verlet integration is only $O(\Delta t^4)$, which implies a more accurate result. Fig. 5 shows eight frames of a typical flattening process.

4 Experimental Results

We have conducted experiments on a set of 52 images captured from books and brochures with arbitrary warpings to test the efficiency and stability of the proposed method. The character and word precision on the restored images are computed and compared with those on the original images. Furthermore, we also compared our method with Brown and Seales' method using a multifolded page image. In addition, examples on extremely crumpled pages are also shown to discuss the breakdown points of the current method. In all the experiments, the 3D models are acquired using Minolta Vivid 900 by using a mid lens with a measuring distance of about 0.9 m. The simulation process runs on a Pentium III 996 MHz PC (512 MB RAM).

4.1 Experiment I: Real Images of Books and Brochures

In this experiment, we evaluate the restoration results based on their OCR performance. The images used here are captured from various sources such as thick bound books, advertising brochures, and posters, etc. Therefore, they span a wide range of warping distortions such as folding, rolling curl, binding curl, rumples, etc. A total of 52 images are examined. Fig. 6 shows eight selected images with different warping curvatures and their restored images. For better visualization and OCR performance, shadings in the restored images are removed using an inpainting-based algorithm [28]. Graphical contents as shown in Figs. 6b, 6e, and 6h are also restored back to their normal size and layout. This demonstrates that our method is content independent and can be applied to a large set of documents including those fragile art materials or newspapers to be digitized in many digital libraries. Besides simple warpings such as rolling curls, binding curls, and multifolds, we also tested on randomly crumpled pages with different degrees of rumples and creases. Fig. 6 f_1 shows a crumpled document image with large creases but less fine rumples. In order to represent the fine rumples more accurately, we use a relatively dense mesh with about 5,000 points. Fig. 6 f_2 shows the restored image.

Next, we feed both the original images and the restored images into OmniPage OCR for recognition. The word and character precisions are computed and compared in Table 1. In general, for a set of 52 images with a total of 12,000 words, the average improvements on word precision and character precision are 18.8 percent and 16.5 percent, respectively. Due to the low resolution of the input images captured by the laser range scanner (640×480), the OCR precision tends to be low for some document images with smaller font. This can be further enhanced through a super-resolution algorithm such as [29]. It is also possible that the 3D scanners could incorporate a higher resolution 2D imaging function in the future.

4.2 Experiment II: Comparisons with Brown and Seales' Method

In comparison with the geometric restoration framework proposed by Brown and Seales [2], [3], our method has made two key improvements in terms of stability and efficiency. First, our framework is able to handle irregular triangular meshes, which better represent some complex shapes compared to quad meshes when they are both sparse. The use of sparse meshes also helps to reduce the simulation time tremendously. Second, by using a stick constraint model, we eliminate the potential instability problems caused by the manually adjusted parameters in the mass-spring model. Fig. 7 compares the results of a multifolded paper restored using our stick model and the existing mass-spring model. Fig. 7c shows the restored mesh using the mass-spring model with the Brown and Seales' original implementation, which is overall satisfactory except some minor errors at the right boundary where some mesh points overlapped with each other during flattening and caused some distortions in the restored image. More overlappings are observed at the folds when the mesh is denser and the spring constants are less carefully tuned. It also takes a longer time to reach a stable state during flattening. Our method is able to deal with denser meshes without self-folding problems attributing to the bending resistance and the drag force as described earlier. Figs. 7d and 7f show the restored mesh and image, respectively. In terms of speed, the existing method takes about 50 seconds to converge while our method only needs 3.13 seconds with two meshes containing both 999 points. To show the efficiency of our method, we evaluated the simulation time on the eight images shown in Fig. 6 and compared with that derived from the existing method. All the above experiments were run on an Intel Pentium III 996MHz PC (512 MB RAM). Time step used for Verlet integration is 0.1 seconds. The comparison shows that our method enhances the efficiency significantly (see Table 2).

For numerical simulation, stability means the simulation is convergent under the condition of the adopted time step. Therefore,

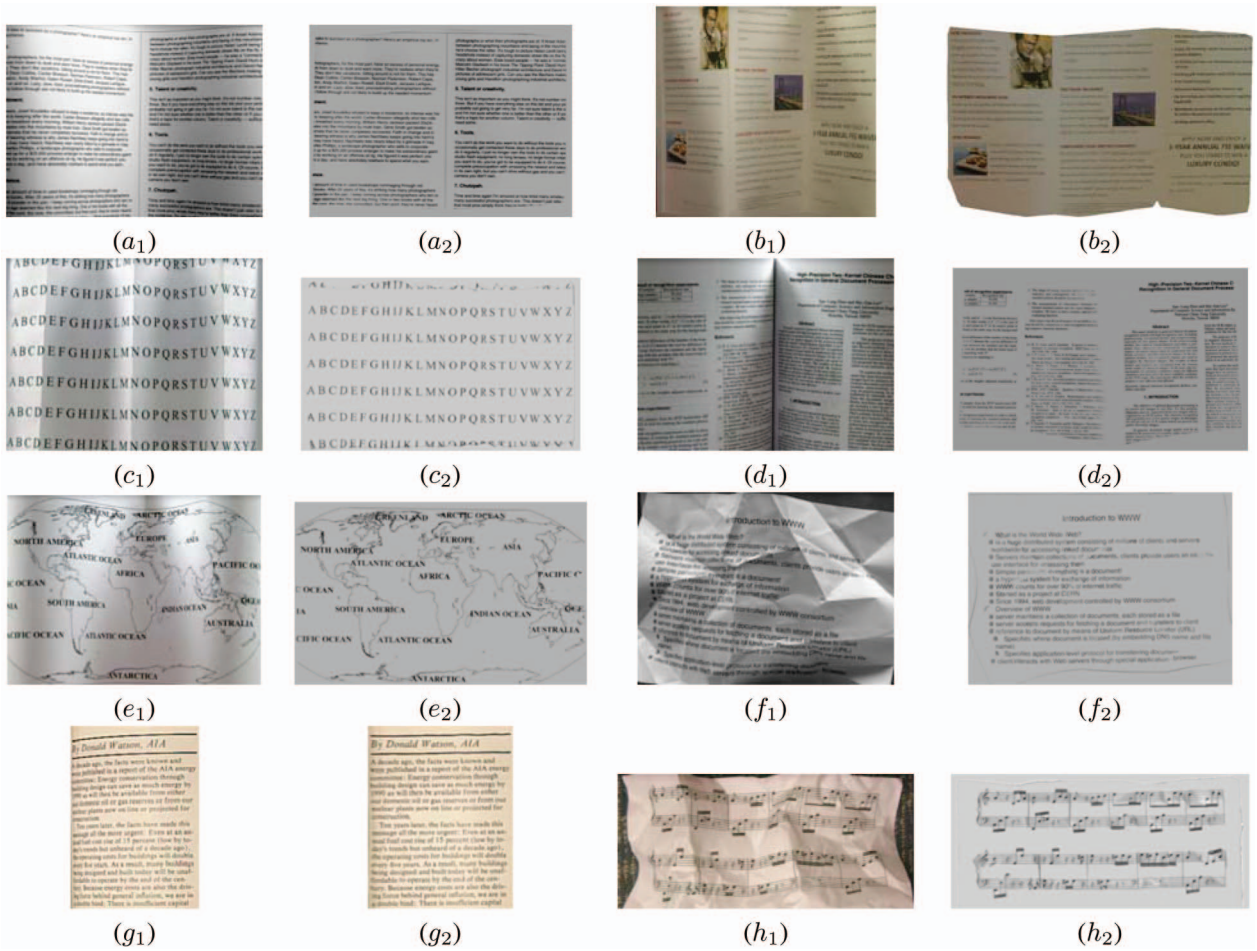

Fig. 6. (a_1)-(h_1) Original distorted 2D image. (a_2)-(h_2) Geometrically restored images with shadings removed.

TABLE 1
Comparisons of OCR Results on Images Shown in Fig. 6

Images	OCR Word Precision		OCR Character Precision	
	Original Image	Restored Image	Original Image	Restored Image
Fig. 6(a)	78.6%	92.3%	83.4%	95.6%
Fig. 6(b)	64.1%	82.6%	80.5%	91.8%
Fig. 6(c)	88.7%	100%	88.7%	100%
Fig. 6(d)	64.9%	85.4%	81.2%	92.4%
Fig. 6(f)	51.4%	94.3%	57.8%	97.7%
Fig. 6(g)	83.3%	93.8%	86.4%	96.8%

it can be measured quantitatively with the maximal time step a method can take without being divergent. For all explicit numerical simulation approaches such as Euler and Runge-Kutta, they are conditionally stable, i.e., there is a maximum time step above which these approaches will become divergent. Normally Euler and Runge-Kutta approaches need small time steps to maintain stability. The use of Verlet integration in our method allows us to use a large time step without divergence. Meanwhile, the simulation needs very few iterations to reach the equilibrium state and thus achieves computational efficiency as well. On average, the maximum time step our method can adopt is more than 10 times larger than that of Brown and Seales' method. In particular, for the example shown in Fig. 7, the maximum time step our method can take is 0.32 seconds in contrast to the 0.02 seconds in the existing method under our simulation.

4.3 Experiment III: Crumpled Pages

Furthermore, Fig. 8 shows the results of a crumpled page with more severe rumples. Although there are many fine wrinkles across the page, no interiors are occluded by other parts of the surface. Therefore, the image can still be restored nicely without loss of content as shown in Fig. 8b. However, if we further increase the level of distortion as shown in Fig. 8c, some parts of the surface will be occluded from the scanner so that the structure and the image are not correctly captured. Therefore, the restored image will lose information at the occluded parts such as the disconnected square edges as shown in Fig. 8d.

4.4 Discussion

Besides the good performance of the presented system, some limitations still remain. One is the low resolution of the 2D images

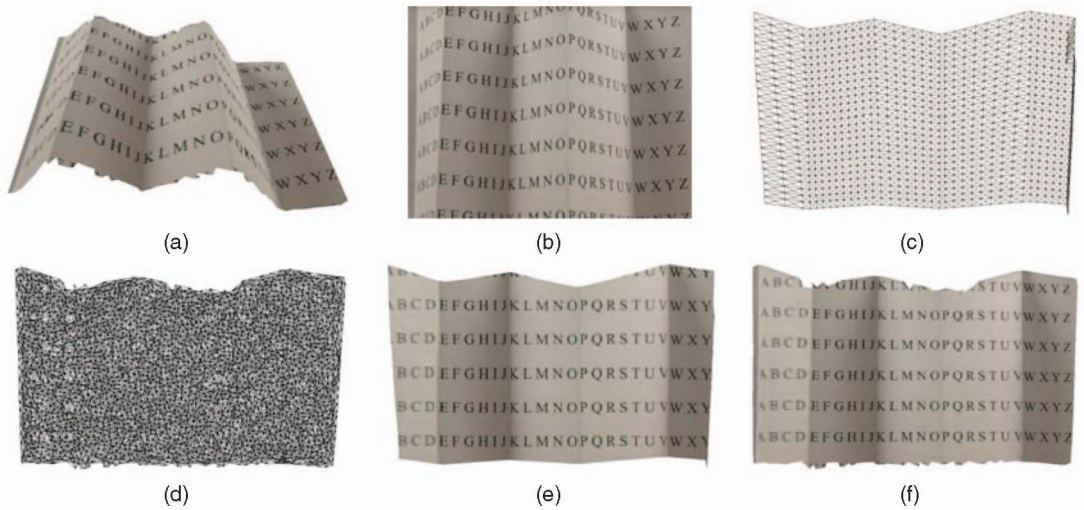

Fig. 7. (a) A multifolded page. (b) Original 2D warped image. (c) Restored mesh using Brown and Seales' method. (d) Restored mesh using our method. (e) Restored image using Brown and Seales' method. (f) Restored image using our triangular mesh with stick model.

TABLE 2
Evaluation of Speed on Images Shown in Fig. 6

Images	Our Method				Brown and Seales'
	Mesh Points	Add BR (s)	Flattening (s)	Total Time (s)	Total Time (s)
Fig. 6(a)	1087	0.25	5.37	5.62	52.35
Fig. 6(b)	1264	0.27	9.25	9.52	64.36
Fig. 6(c)	1287	0.28	9.68	9.96	66.42
Fig. 6(d)	1154	0.26	5.56	5.82	56.67
Fig. 6(e)	1181	0.26	7.25	7.51	63.78
Fig. 6(f)	4986	0.34	10.35	10.69	72.43
Fig. 6(g)	1446	0.29	7.87	8.16	62.54
Fig. 6(h)	4732	0.33	10.13	10.46	71.07

captured by the laser range scanner. This turns up to be a disadvantage compared to those normal digital cameras which can easily capture high resolution images of 10 megapixels or even higher. However, since the 3D scanner hardware is upgrading at a fast pace, it is not difficult to incorporate high-resolution 2D imaging functions in the future. On the other hand, we can try to capture multiple low-resolution images to construct a single high resolution image using some existing methods [15], [30]. Also, to ensure that the object being captured is visible to the laser range scanner, its surface must reflect red laser light. Therefore, if the surface is colored as purely black, blue, or green, the shape at the corresponding part will not be captured. This also applies to some highly shiny surfaces due to specular reflection. Nevertheless, if such portions are small such as black colored texts, they still can be captured without much problem. Even if there are unidentified parts, the holes can be filled at a postprocessing phase. Furthermore, if the geometric distortions are too complicated like a severely crumpled page with many hidden edges such as the example shown in Fig. 8c, the restored image will have some missing portions due to the occlusions. This is an inevitable problem and can possibly be resolved through 3D inpainting. Lastly, considering the price of the current laser range scanners, it might be a high cost for some normal users. A remedy to this problem is to apply some shape recovery techniques such as Shape-from-X to reconstruct the shape from the given 2D image and then apply the flattening procedure. However, to recover an accurate shape from a single 2D image is also a challenging problem which needs to be further studied.

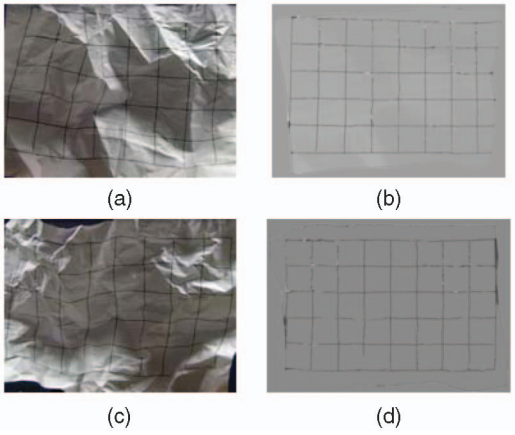

Fig. 8. (a) A crumpled page with more severe rumples. (b) Geometrically restored image with shadings removed. (c) A severely crumpled image with occlusions. (d) Geometrically restored image with shadings removed.

5 Conclusion

We have presented a physically-based restoration method with improved efficiency and stability for correcting geometric distortions in arbitrarily warped document images. In the proposed method, the 3D model and the 2D texture are acquired through a

laser range scanner. A mesh refinement step is done prior to the flattening process to subsample the original dense mesh and add a bending resistance to increase the rigidity of the mesh. The refined mesh is then represented as a particle system and flattened using Verlet integration. During the simulation, two external forces, gravitational force and a drag force, are exerted on the particles to drive them to the $z = 0$ plane. Meanwhile, stick constraint and plane collision constraint are satisfied, which retains the rigidity of the mesh and guarantees the convergence of the system. In comparison with Brown and Seales' method, the current method achieves considerable improvements in terms of stability and efficiency. As physically-based restoration approaches are content independent and, thus, can handle either textual documents or graphical documents or a mixture of the two, our improved method is especially helpful to those precious ancient manuscripts and art materials in digital libraries. With the fast advancing imaging hardware, it is very likely that the future cell phones or PDAs will be empowered with 3D imaging functions. This will make our method a useful application in current prevalent digital devices.

ACKNOWLEDGMENTS

This project is supported in part by Agency for Science, Technology and Research (A*STAR) under grant no. 042-101-0085 and National University of Singapore under URC grant no. R252-000-202-112.

REFERENCES

[1] M. Pilu, "Undoing Page Curl Distortion Using Applicable Surfaces," *Computer Vision and Pattern Recognition*, vol. 1, pp. 67-72, 2001.
[2] M. Brown and W. Seales, "Document Restoration Using 3D Shape: A General Deskewing Algorithm for Arbitrarily Warped Documents," *Proc. Int'l Conf. Computer Vision*, vol. 2, pp. 367-374, 2001.
[3] M. Brown and W. Seales, "Image Restoration of Arbitrarily Warped Documents," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 26, no. 10, pp. 1295-1306, Oct. 2004.
[4] M. Brown and C. Pisula, "Conformal Deskewing of Non-Planar Documents," *Proc. IEEE CS Conf. Computer Vision and Pattern Recognition*, vol. 1, pp. 998-1004, 2005.
[5] J. Liang, D. DeMenthon, and D. Doermann, "Camera-Based Analysis of Text and Documents: A Survey," *Int'l J. Document Analysis and Recognition*, vol. 7, nos. 2-3, pp. 83-104, 2005.
[6] Y.Y. Tang and C.Y. Suen, "Image Transformation Approach to Nonlinear Shape Restoration," *IEEE Trans. Systems, Man, and Cybernetics*, vol. 1, no. 23, pp. 155-171, 1993.
[7] Z. Zhang and C.L. Tan, "Straightening Warped Text Lines Using Polynomial Regression," *Proc. Int'l Conf. Image Processing*, pp. 977-980, 2002.
[8] Y.C. Tsoi and M.S. Brown, "Geometric and Shading Correction for Images of Printed Materials—A Unified Approach Using Boundary," *Computer Vision and Pattern Recognition*, vol. 1, pp. 240-246, 2004.
[9] M. Sun, R. Yang, L. Yun, G. Landon, B. Seales, and M. Brown, "Geometric and Photometric Restoration of Distorted Documents," *Proc. 10th IEEE Int'l Conf. Computer Vision*, vol. 2, pp. 17-21, 2005.
[10] K.B. Chua, L. Zhang, Y. Zhang, and C.L. Tan, "A Fast and Stable Approach for Restoration of Warped Document Images," *Proc. IAPR Int'l Conf. Document Analysis and Recognition*, no. 1, pp. 384-388, 2005.
[11] T. Amano, T. Abe, T. Iyoda, O. Nishikawa, and Y. Sato, "Camera-Based Document Image Mosaicing," *Proc. SPIE*, vol. 4669, pp. 250-258, 2002.
[12] W. Seales and Y. Lin, "Digital Restoration Using Volumetric Scanning," *Proc. Joint ACM/IEEE Conf. Digital Libraries*, vol. 1, pp. 117-124, 2004.
[13] Y. Lin and W. Seales, "Opaque Document Imaging: Building Images of Inaccessible Texts," *Proc. 10th IEEE Int'l Conf. Computer Vision*, vol. 1, pp. 662-669, 2005.
[14] A. Yamashita, A. Kawarago, T. Kaneko, and K.T. Miura, "Shape Reconstruction and Image Restoration for Non-Flat Surface of Document with a Stereo Vision System," *Proc. IEEE Int'l Conf. Pattern Recognition*, pp. 482-485, 2004.
[15] A. Iketani, T. Sato, and S. Ikeda, "Video Mosaicing for Curved Documents Based on Structure from Motion," *Proc. 18th Int'l Conf. Pattern Recognition*, vol. 1, pp. 391-396, 2006.
[16] "3D Digitizers—Non-Contact Laser Range Scanner,"Ramsey, USA, Konica Minolta, http://www.minolta3d.com, 2007.
[17] A. Witkin, "An Introduction to Physically Based Modeling: Particle System Dynamics," *Proc. ACM SIGGRAPH*, 1997.
[18] D. Baraff, "An Introduction to Physically Based Modeling: Rigid Body Simulation I—Unconstrained Rigid Body Dynamics," *Proc. ACM SIGGRAPH Tutorial Notes*, 1997.
[19] G. Oliveira, "Exploring Spring Models," *Game Developer*, 2001.
[20] X. Provot, "Deformation Constraints in a Mass-Spring Model to Describe Rigid Cloth Behaviour," *Graphics Interface*, pp. 155-174, 1995.
[21] T. Jakobsen, "Advanced Character Physics," *Game Developer*, 2003.
[22] K.J. Choi and H.S. Ko, "Extending the Immediate Buckling Model to Triangular Meshes for Simulating Complex Clothes," *Proc. EUROGRAPHICS*, pp. 187-191, 2003.
[23] P. Volino and N. Magnenat-Thalmann, "Efficient Self-Collision Detection on Smoothly Discretized Surface Animations Using Geometrical Shape Regularity," *Proc. EUROGRAPHICS*, no. 13, pp. 155-166, 1994.
[24] P. Volino and N. Magnenat-Thalmann, "Implementing Fast Cloth Simulation with Collision Response," *Proc. Computer Graphics Int'l*, pp. 257-268, 2000.
[25] X. Provot, "Collision and Self-Collision Detection Handling in Cloth Model Dedicated to Design Garments," *Graphics Interface*, pp. 177-189, 1997.
[26] A. Witkin, "Physically Based Modeling - Particle System Dynamics," *Proc. ACM SIGGRAPH Course Notes*, 2001.
[27] W.H. Press, B.P. Fannery, S.A. Teukolsky, and W.T. Verrerling, *Numerical Recipes: The Art of Scientific Computing*. Cambridge Univ. Press, 1986.
[28] L. Zhang, A. Yip, and C. Tan, "Photometric and Geometric Restoration of Document Images Using Inpainting and Shape-from-Shading," *Proc. 22nd Conf. Artificial Intelligence*, 2007.
[29] S. Farsiu, D. Robinson, M. Elad, and P. Milanfar, "Fast and Robust Super-Resolution," *Proc. IEEE Int'l Conf. Image Processing*, pp. 291-294, 2003.
[30] J. Liang, D. DeMenthon, and D. Doermann, "Camera-Based Document Image Mosaicing," *Proc. Int'l Conf. Pattern Recognition*, no. 2, pp. 476-479, 2006.

▷ **For more information on this or any other computing topic, please visit our Digital Library at** www.computer.org/publications/dlib.