

SIT789 - Robotics Computer Vision And Speech Processing

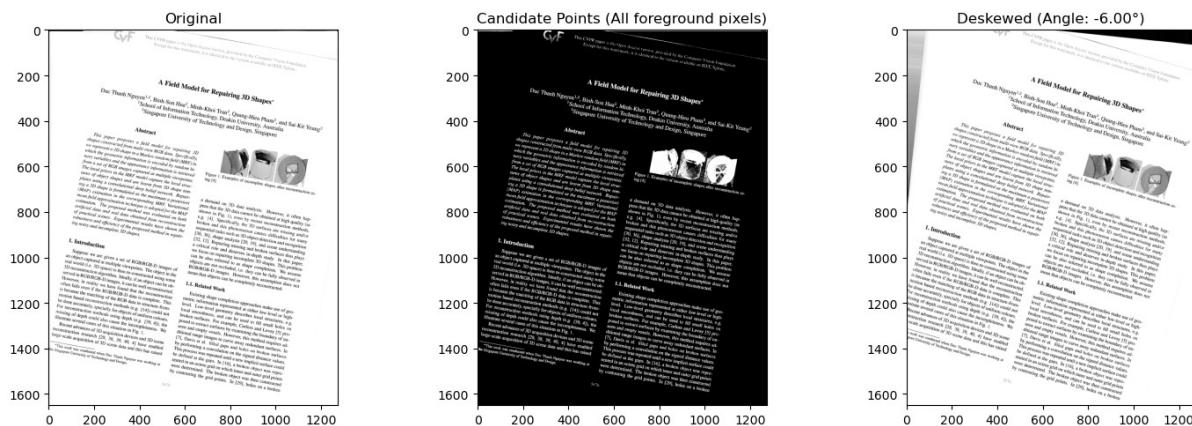
Distinction Task 2.3: Document analysis and recognition

Introduction

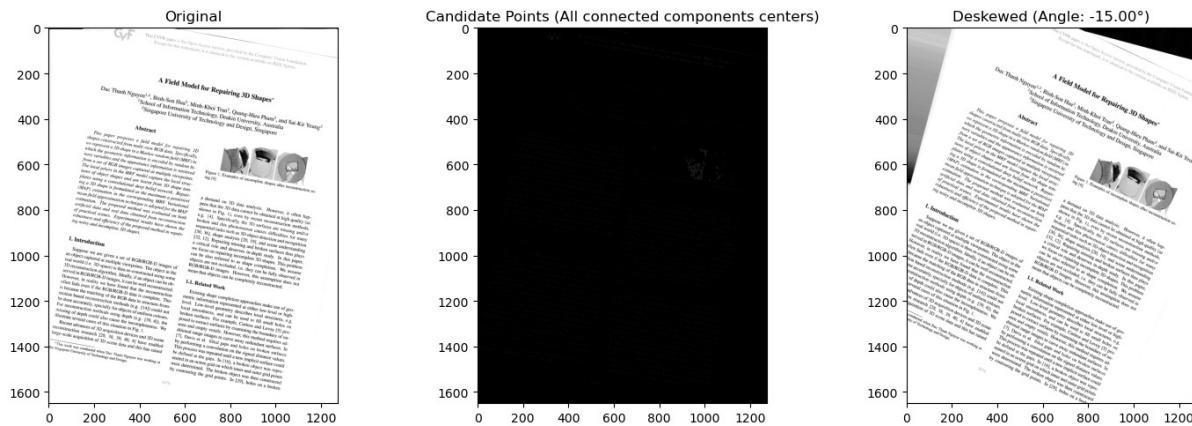
Document analysis and recognition play a vital role in transforming document images into text-based formats. This transformation facilitates various functionalities such as content retrieval, searching, indexing, and document compression.

1.

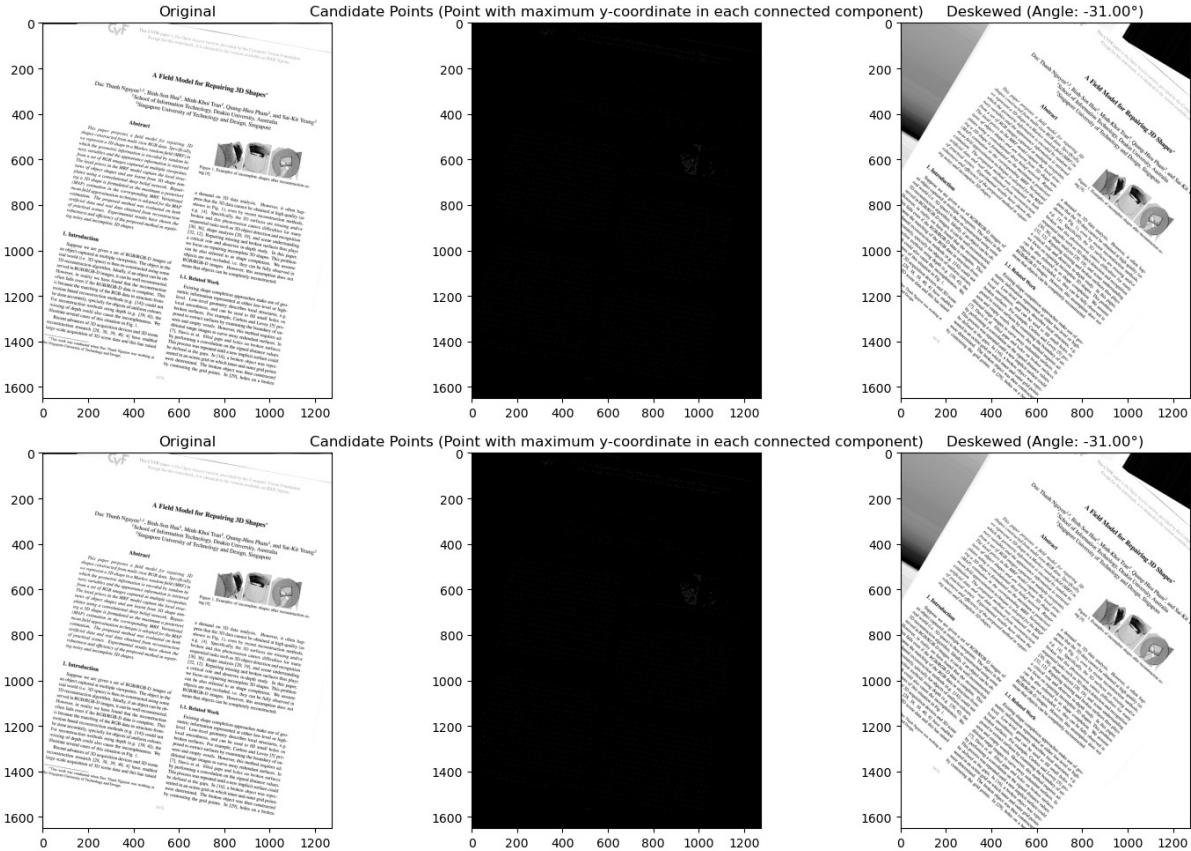
(a) All foreground pixels strategy



(b) All connected components centers:



(c) Point with maximum y-coordinate in each connected component



2. De-skewing results:

Strategy: A) All foreground pixels strategy

Density threshold: 5

A Field Model for Repairing 3D Shapes*

Duc Thanh Nguyen^{1,2}, Binh-Son Hua², Minh-Khoi Tran², Quang-Hieu Pham², and Sai-Kit Yeung²

¹School of Information Technology, Deakin University, Australia
²Singapore University of Technology and Design, Singapore

Abstract

This paper proposes a field model for repairing 3D shapes constructed from multi-view RGB data. Specifically, we represent a 3D shape in a Markov random field (MRF) in which the geometric information is encoded by random binary variables and the appearance information is retrieved from a set of RGB images captured at multiple viewpoints. The local priors in the MRF model capture the local structures of object shapes and are learnt from 3D shape templates using a convolutional deep belief network. Repairing a 3D shape is formulated as the maximum a posteriori (MAP) estimation in the corresponding MRF. Variational mean field approximation technique is adopted for the MAP estimation. The proposed method was evaluated on both artificial data and real data obtained from reconstruction of practical scenes. Experimental results have shown the robustness and efficiency of the proposed method in repairing noisy and incomplete 3D shapes.

1. Introduction

Suppose we are given a set of RGB/RGB-D images of an object captured at multiple viewpoints. The object in the real world (i.e. 3D space) is then re-constructed using some 3D reconstruction algorithm. Ideally, if an object can be observed in RGB/RGB-D images, it can be well reconstructed. However, in reality we have found that the reconstruction often fails even if the RGB/RGB-D data is complete. This is because the matching of the RGB data in structure-from-motion based reconstruction methods (e.g. [14]) could not be done accurately, especially for objects of uniform colours. For reconstruction methods using depth (e.g. [39, 4]), the missing of depth could also cause the incompleteness. We illustrate several cases of this situation in Fig. 1. Recent advances of 3D acquisition devices and 3D scene construction research [28, 38, 39, 40, 4] have enabled re-scale acquisition of 3D scene data and this has raised

This work was conducted when Duc Thanh Nguyen was working at Singapore University of Technology and Design.

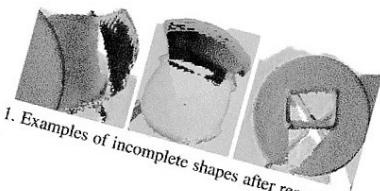


Figure 1. Examples of incomplete shapes after reconstruction using [4].

a demand on 3D data analysis. However, it often happens that the 3D data cannot be obtained at high quality (as shown in Fig. 1), even by recent reconstruction methods, e.g. [4]. Specifically, the 3D surfaces are missing and/or broken and this phenomenon causes difficulties for many sequential tasks such as 3D object detection and recognition [30, 36], shape analysis [20, 19], and scene understanding [32, 12]. Repairing missing and broken surfaces thus plays a critical role and deserves in-depth study. In this paper, we focus on repairing incomplete 3D shapes. This problem can be also referred to as shape completion. We assume objects are not occluded, i.e. they can be fully observed in RGB/RGB-D images. However, this assumption does not mean that objects can be completely reconstructed.

1.1. Related Work

Existing shape completion approaches make use of geometric information represented at either low-level or high-level. Low-level geometry describes local structures, e.g. local smoothness, and can be used to fill small holes on broken surfaces. For example, Curless and Levoy [5] proposed to extract surfaces by examining the boundary of un-seen and empty voxels. However, this method requires additional range images to carve away redundant surfaces. In [7], Davis et al. filled gaps and holes on broken surfaces by performing a convolution on the signed distance values. This process was repeated until a new implicit surface could be defined at the gaps. In [16], a broken object was represented in an octree grid on which inner and outer grid points were determined. The broken object was then constructed by contouring the grid points. In [29], holes on a broken

A Field Model for Repairing 3D Shapes*

Duc Thanh Nguyen^{1,2}, Binh-Son Hua², Minh-Khoi Tran², Quang-Hieu Pham², and Sai-Kit Yeung²

¹School of Information Technology, Deakin University, Australia
²Singapore University of Technology and Design, Singapore

Abstract

This paper proposes a field model for repairing 3D shapes constructed from multi-view RGB data. Specifically, we represent a 3D shape in a Markov random field (MRF) in which the geometric information is encoded by random binary variables and the appearance information is retrieved from a set of RGB images captured at multiple viewpoints. The local priors in the MRF model capture the local structures of object shapes and are learnt from 3D shape templates using a convolutional deep belief network. Repairing a 3D shape is formulated as the maximum a posteriori (MAP) estimation in the corresponding MRF. Variational mean-field approximation technique is adopted for the MAP estimation. The proposed method was evaluated on both artificial data and real data obtained from reconstruction of practical scenes. Experimental results have shown the robustness and efficiency of the proposed method in repairing noisy and incomplete 3D shapes.

1. Introduction

Suppose we are given a set of RGB/RGB-D images of an object captured at multiple viewpoints. The object in the real world (i.e. 3D space) is then re-constructed using some 3D reconstruction algorithm. Ideally, if an object can be observed in RGB/RGB-D images, it can be well reconstructed. However, in reality we have found that the reconstruction often fails even if the RGB/RGB-D data is complete. This is because the matching of the RGB data in structure-from-motion based reconstruction methods (e.g. [14]) could not be done accurately, specially for objects of uniform colours. For reconstruction methods using depth (e.g. [39, 4]), the missing of depth could also cause the incompleteness. We illustrate several cases of this situation in Fig. 1. Recent advances of 3D acquisition devices and 3D scene reconstruction research [28, 38, 39, 40, 4] have enabled large-scale acquisition of 3D scene data and this has raised

* work was conducted when Duc Thanh Nguyen was working at Singapore University of Technology and Design.

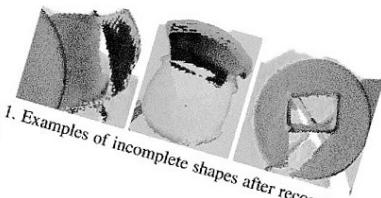


Figure 1. Examples of incomplete shapes after reconstruction using [4].

a demand on 3D data analysis. However, it often happens that the 3D data cannot be obtained at high quality (as shown in Fig. 1), even by recent reconstruction methods, e.g. [4]. Specifically, the 3D surfaces are missing and/or broken and this phenomenon causes difficulties for many sequential tasks such as 3D object detection and recognition [30, 36], shape analysis [20, 19], and scene understanding [32, 12]. Repairing missing and broken surfaces thus plays a critical role and deserves in-depth study. In this paper, we focus on repairing incomplete 3D shapes. This problem can be also referred to as shape completion. We assume objects are not occluded, i.e. they can be fully observed in RGB/RGB-D images. However, this assumption does not mean that objects can be completely reconstructed.

1.1. Related Work

Existing shape completion approaches make use of geometric information represented at either low-level or high-level. Low-level geometry describes local structures, e.g. local smoothness, and can be used to fill small holes on broken surfaces. For example, Curless and Levoy [5] proposed to extract surfaces by examining the boundary of un-seen and empty voxels. However, this method requires additional range images to carve away redundant surfaces. In [7], Davis et al. filled gaps and holes on broken surfaces by performing a convolution on the signed distance values. This process was repeated until a new implicit surface could be defined at the gaps. In [16], a broken object was represented in an octree grid on which inner and outer grid points were determined. The broken object was then constructed by contouring the grid points. In [29], holes on a broken

A Field Model for Repairing 3D Shapes*

Duc Thanh Nguyen^{1,2}, Binh-Son Hua², Minh-Khoi Tran², Quang-Hieu Pham², and Sai-Kit Yeung²

¹School of Information Technology, Deakin University, Australia
²Singapore University of Technology and Design, Singapore

Abstract

This paper proposes a field model for repairing 3D shapes constructed from multi-view RGB data. Specifically, we represent a 3D shape in a Markov random field (MRF) in which the geometric information is encoded by random binary variables and the appearance information is retrieved from a set of RGB images captured at multiple viewpoints. The local priors in the MRF model capture the local structures of object shapes and are learnt from 3D shape templates using a convolutional deep belief network. Repairing a 3D shape is formulated as the maximum a posteriori (MAP) estimation in the corresponding MRF. Variational mean-field approximation technique is adopted for the MAP estimation. The proposed method was evaluated on both artificial data and real data obtained from reconstruction of practical scenes. Experimental results have shown the robustness and efficiency of the proposed method in repairing noisy and incomplete 3D shapes.

1. Introduction

Suppose we are given a set of RGB/RGB-D images of an object captured at multiple viewpoints. The object in the real world (i.e. 3D space) is then re-constructed using some 3D reconstruction algorithm. Ideally, if an object can be observed in RGB/RGB-D images, it can be well reconstructed. However, in reality we have found that the reconstruction often fails even if the RGB/RGB-D data is complete. This is because the matching of the RGB data in structure-from-motion based reconstruction methods (e.g. [14]) could not be done accurately, specially for objects of uniform colours. For reconstruction methods using depth (e.g. [39, 4]), the missing of depth could also cause the incompleteness. We illustrate several cases of this situation in Fig. 1. Recent advances of 3D acquisition devices and 3D scene construction research [28, 38, 39, 40, 4] have enabled large-scale acquisition of 3D scene data and this has raised

* work was conducted when Duc Thanh Nguyen was working at Singapore University of Technology and Design.

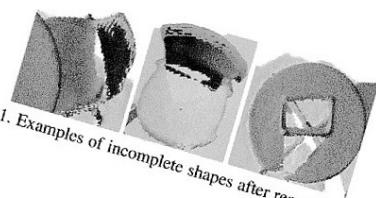


Figure 1. Examples of incomplete shapes after reconstruction using [4].

a demand on 3D data analysis. However, it often happens that the 3D data cannot be obtained at high quality (as shown in Fig. 1), even by recent reconstruction methods, e.g. [4]. Specifically, the 3D surfaces are missing and/or broken and this phenomenon causes difficulties for many sequential tasks such as 3D object detection and recognition [30, 36], shape analysis [20, 19], and scene understanding [32, 12]. Repairing missing and broken surfaces thus plays a critical role and deserves in-depth study. In this paper, we focus on repairing incomplete 3D shapes. This problem can be also referred to as shape completion. We assume objects are not occluded, i.e. they can be fully observed in RGB/RGB-D images. However, this assumption does not mean that objects can be completely reconstructed.

1.1. Related Work

Existing shape completion approaches make use of geometric information represented at either low-level or high-level. Low-level geometry describes local structures, e.g. local smoothness, and can be used to fill small holes on broken surfaces. For example, Curless and Levoy [5] proposed to extract surfaces by examining the boundary of unseen and empty voxels. However, this method requires additional range images to carve away redundant surfaces. In [7], Davis et al. filled gaps and holes on broken surfaces by performing a convolution on the signed distance values. This process was repeated until a new implicit surface could be defined at the gaps. In [16], a broken object was represented in an octree grid on which inner and outer grid points were determined. The broken object was then constructed by contouring the grid points. In [29], holes on a broken

Density threshold: 20



This CVPR paper is the Open Access version, provided by the Computer Vision Foundation.
Except for this watermark, it is identical to the version available on IEEE Xplore.

A Field Model for Repairing 3D Shapes*

Duc Thanh Nguyen^{1,2}, Binh-Son Hua², Minh-Khoi Tran², Quang-Hieu Pham², and Sai-Kit Yeung²

¹School of Information Technology, Deakin University, Australia
²Singapore University of Technology and Design, Singapore

Abstract

This paper proposes a field model for repairing 3D shapes constructed from multi-view RGB data. Specifically, we represent a 3D shape in a Markov random field (MRF) in which the geometric information is encoded by random binary variables and the appearance information is retrieved from a set of RGB images captured at multiple viewpoints. The local priors in the MRF model capture the local structures of object shapes and are learnt from 3D shape templates using a convolutional deep belief network. Repairing a 3D shape is formulated as the maximum a posteriori (MAP) estimation in the corresponding MRF. Variational mean field approximation technique is adopted for the MAP estimation. The proposed method was evaluated on both artificial data and real data obtained from reconstruction of practical scenes. Experimental results have shown the robustness and efficiency of the proposed method in repairing noisy and incomplete 3D shapes.

1. Introduction

Suppose we are given a set of RGB/RGB-D images of an object captured at multiple viewpoints. The object in the real world (i.e. 3D space) is then re-constructed using some 3D reconstruction algorithm. Ideally, if an object can be observed in RGB/RGB-D images, it can be well reconstructed. However, in reality we have found that the reconstruction often fails even if the RGB/RGB-D data is complete. This is because the matching of the RGB data in structure-from-motion based reconstruction methods (e.g. [14]) could not be done accurately, specially for objects of uniform colours. For reconstruction methods using depth (e.g. [39, 4]), the missing of depth could also cause the incompleteness. We illustrate several cases of this situation in Fig. 1. Recent advances of 3D acquisition devices and 3D scene reconstruction research [28, 38, 39, 40, 4] have enabled large-scale acquisition of 3D scene data and this has raised

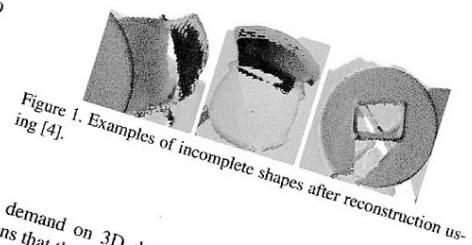


Figure 1. Examples of incomplete shapes after reconstruction using [4].

a demand on 3D data analysis. However, it often happens that the 3D data cannot be obtained at high quality (as shown in Fig. 1), even by recent reconstruction methods, e.g. [4]. Specifically, the 3D surfaces are missing and/or broken and this phenomenon causes difficulties for many sequential tasks such as 3D object detection and recognition [30, 36], shape analysis [20, 19], and scene understanding [32, 12]. Repairing missing and broken surfaces thus plays a critical role and deserves in-depth study. In this paper, we focus on repairing incomplete 3D shapes. This problem can be also referred to as shape completion. We assume objects are not occluded, i.e. they can be fully observed in RGB/RGB-D images. However, this assumption does not mean that objects can be completely reconstructed.

1.1. Related Work

Existing shape completion approaches make use of geometric information represented at either low-level or high-level. Low-level geometry describes local structures, e.g. local smoothness, and can be used to fill small holes on broken surfaces. For example, Curless and Levoy [5] proposed to extract surfaces by examining the boundary of unseen and empty voxels. However, this method requires additional range images to carve away redundant surfaces. In [7], Davis et al. filled gaps and holes on broken surfaces by performing a convolution on the signed distance values. This process was repeated until a new implicit surface could be defined at the gaps. In [16], a broken object was represented in an octree grid on which inner and outer grid points were determined. The broken object was then constructed by contouring the grid points. In [29], holes on a broken

work was conducted when Duc Thanh Nguyen was working at Nanyang Technological University of Technology and Design.

Strategy: B) All connected components' centers

Density threshold: 5

This CVPR paper is the Open Access version, provided by the Computer Vision Foundation.
Except for this watermark, it is identical to the version available on IEEE Xplore.

A Field Model for Repairing 3D Shapes*

Duc Thanh Nguyen^{1,2}, Binh-Son Hua², Minh-Khoi Tran², Quang-Hieu Pham², and Sai-Kit Yeung²

¹School of Information Technology, Deakin University, Australia
²Singapore University of Technology and Design, Singapore

Abstract

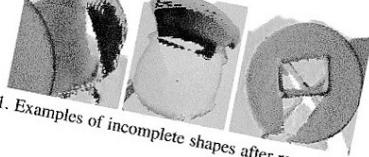
This paper proposes a field model for repairing 3D shapes constructed from multi-view RGB data. Specifically, we represent a 3D shape in a Markov random field (MRF) in which the geometric information is encoded by random binary variables and the appearance information is retrieved from a set of RGB images captured at multiple viewpoints. The local priors in the MRF model capture the local structures of object shapes and are learnt from 3D shape templates using a convolutional deep belief network. Repairing a 3D shape is formulated as the maximum a posteriori (MAP) estimation in the corresponding MRF. Variational mean field approximation technique is adopted for the MAP estimation. The proposed method was evaluated on both artificial data and real data obtained from reconstruction of practical scenes. Experimental results have shown the robustness and efficiency of the proposed method in repairing noisy and incomplete 3D shapes.

1. Introduction

Suppose we are given a set of RGB/RGB-D images of an object captured at multiple viewpoints. The object in the real world (i.e. 3D space) is then re-constructed using some 3D reconstruction algorithm. Ideally, if an object can be observed in RGB/RGB-D images, it can be well reconstructed. However, in reality we have found that the reconstruction often fails even if the RGB/RGB-D data is complete. This is because the matching of the RGB data in structure-from-motion based reconstruction methods (e.g. [14]) could not be done accurately, especially for objects of uniform colours. For reconstruction methods using depth (e.g. [39, 4]), the missing of depth could also cause the incompleteness. We illustrate several cases of this situation in Fig. 1.

Recent advances of 3D acquisition devices and 3D scene reconstruction research [28, 38, 39, 40, 4] have enabled large-scale acquisition of 3D scene data and this has raised

Figure 1. Examples of incomplete shapes after reconstruction using [4].



a demand on 3D data analysis. However, it often happens that the 3D data cannot be obtained at high quality (as shown in Fig. 1), even by recent reconstruction methods, e.g. [4]. Specifically, the 3D surfaces are missing and/or broken and this phenomenon causes difficulties for many sequential tasks such as 3D object detection and recognition [30, 36], shape analysis [20, 19], and scene understanding [32, 12]. Repairing missing and broken surfaces thus plays a critical role and deserves in-depth study. In this paper, we focus on repairing incomplete 3D shapes. This problem can be also referred to as shape completion. We assume objects are not occluded, i.e. they can be fully observed in RGB/RGB-D images. However, this assumption does not mean that objects can be completely reconstructed.

1.1. Related Work

Existing shape completion approaches make use of geometric information represented at either low-level or high-level. Low-level geometry describes local structures, e.g. local smoothness, and can be used to fill small holes on broken surfaces. For example, Curless and Levoy [5] proposed to extract surfaces by examining the boundary of unseen and empty voxels. However, this method requires additional range images to carve away redundant surfaces. In [7], Davis et al. filled gaps and holes on broken surfaces by performing a convolution on the signed distance values. This process was repeated until a new implicit surface could be defined at the gaps. In [16], a broken object was represented in an octree grid on which inner and outer grid points were determined. The broken object was then constructed by contouring the grid points. In [29], holes on a broken

*This work was conducted when Duc Thanh Nguyen was working at the Singapore University of Technology and Design.

Density threshold: 10

This CVPR paper is the Open Access version, provided by the Computer Vision Foundation. Except for this watermark, it is identical to the version available on IEEE Xplore.

A Field Model for Repairing 3D Shapes*

Duc Thanh Nguyen^{1,2}, Binh-Son Hua², Minh-Khoi Tran², Quang-Hieu Pham², and Sai-Kit Yeung²
¹School of Information Technology, Deakin University, Australia
²Singapore University of Technology and Design, Singapore

Abstract

This paper proposes a field model for repairing 3D shapes constructed from multi-view RGB data. Specifically, we represent a 3D shape in a Markov random field (MRF) in which the geometric information is encoded by random binary variables and the appearance information is retrieved from a set of RGB images captured at multiple viewpoints. The local priors in the MRF model capture the local structures of object shapes and are learnt from 3D shape templates using a convolutional deep belief network. Repairing a 3D shape is formulated as the maximum a posteriori (MAP) estimation in the corresponding MRF. Variational mean field approximation technique is adopted for the MAP estimation. The proposed method was evaluated on both artificial data and real data obtained from reconstruction of practical scenes. Experimental results have shown the robustness and efficiency of the proposed method in repairing noisy and incomplete 3D shapes.

1. Introduction

Suppose we are given a set of RGB/RGB-D images of an object captured at multiple viewpoints. The object in the real world (i.e. 3D space) is then re-constructed using some 3D reconstruction algorithms. Ideally, if an object can be observed in RGB/RGB-D images, it can be well reconstructed. However, in reality we have found that the reconstruction often fails even if the RGB/RGB-D data is complete. This is because the matching of the RGB data (e.g. [14]) could not be done accurately, especially for objects of uniform colours. Several cases of this situation in Fig. 1. The acquisition of 3D acquisition devices and 3D scene reconstruction methods using depth (e.g. [39, 4]) have enabled several cases of this situation in Fig. 1. The acquisition of 3D acquisition devices and 3D scene reconstruction methods using depth (e.g. [39, 4]) have enabled

a demand on 3D data analysis. However, it often happens that the 3D data cannot be obtained at high quality (as shown in Fig. 1), even by recent reconstruction methods, e.g. [4]. Specifically, the 3D surfaces causes difficulties for many broken and missing surfaces thus plays sequential tasks such as 3D object detection and recognition [30, 36]. Repairing missing and broken surfaces thus plays a critical role and deserves in-depth study. In this paper, we focus on repairing incomplete 3D shapes. We assume objects are not occluded, i.e. they can be fully observed in RGB/RGB-D images. Completely reconstructed,

1.1. Related Work

Existing shape completion approaches make use of geometric information represented at either low-level or high-level. Low-level geometry describes local structures, e.g. local smoothness, and can be used to fill small holes on broken surfaces. For example, Curless and Levoy [5] proposed to extract surfaces by examining the boundary of unseen and empty voxels. However, this method requires additional range images to carve away redundant surfaces. This process was repeated until a new implicit surface could be defined at the gaps. In [16], a broken object was represented in an octree grid on which inner and outer grid points were determined. The broken object was then constructed by contouring the grid points. In [29], holes on a bro-

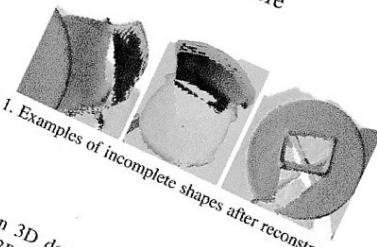
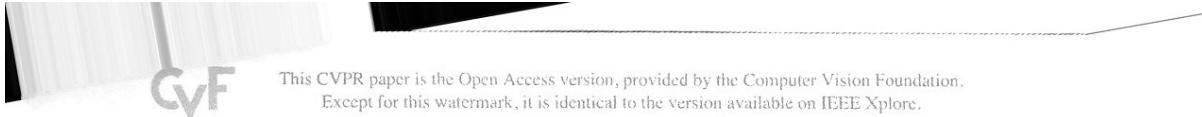


Figure 1. Examples of incomplete shapes after reconstruction using [4].

*conducted when Duc Thanh Nguyen was working at

Density Threshold: 15



A Field Model for Repairing 3D Shapes*

Duc Thanh Nguyen^{1,2}, Binh-Son Hua², Minh-Khoi Tran², Quang-Hieu Pham², and Sai-Kit Yeung²

¹School of Information Technology, Deakin University, Australia

²Singapore University of Technology and Design, Singapore

Abstract

This paper proposes a field model for repairing 3D shapes constructed from multi-view RGB data. Specifically, we represent a 3D shape in a Markov random field (MRF) in which the geometric information is encoded by random binary variables and the appearance information is retrieved from a set of RGB images captured at multiple viewpoints. The local priors in the MRF model capture the local structures of object shapes and are learnt from 3D shape templates using a convolutional deep belief network. Repairing a 3D shape is formulated as the maximum a posteriori (MAP) estimation in the corresponding MRF. Variational mean field approximation technique is adopted for the MAP estimation. The proposed method was evaluated on both artificial data and real data obtained from reconstruction of practical scenes. Experimental results have shown the robustness and efficiency of the proposed method in repairing noisy and incomplete 3D shapes.

1. Introduction

Suppose we are given a set of RGB/RGB-D images of an object captured at multiple viewpoints. The object in the real world (i.e. 3D space) is then re-constructed using some 3D reconstruction algorithm. Ideally, if an object can be observed in RGB/RGB-D images, it can be well reconstructed. However, in reality we have found that the reconstruction often fails even if the RGB/RGB-D data is complete. This is because the matching of the RGB data in structure-from-motion based reconstruction methods (e.g. [14]) could not be done accurately, specially for objects of uniform colours. For reconstruction methods using depth (e.g. [39, 4]), the missing of depth could also cause the incompleteness. We illustrate several cases of this situation in Fig. 1.

Recent advances of 3D acquisition devices and 3D scene reconstruction research [28, 38, 39, 40, 4] have enabled large-scale acquisition of 3D scene data and this has raised

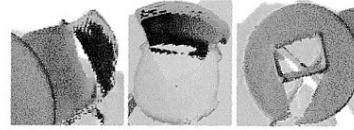


Figure 1. Examples of incomplete shapes after reconstruction using [4].

a demand on 3D data analysis. However, it often happens that the 3D data cannot be obtained at high quality (as shown in Fig. 1), even by recent reconstruction methods, e.g. [4]. Specifically, the 3D surfaces are missing and/or broken and this phenomenon causes difficulties for many sequential tasks such as 3D object detection and recognition [30, 36], shape analysis [20, 19], and scene understanding [32, 12]. Repairing missing and broken surfaces thus plays a critical role and deserves in-depth study. In this paper, we focus on repairing incomplete 3D shapes. This problem can be also referred to as shape completion. We assume objects are not occluded, i.e. they can be fully observed in RGB/RGB-D images. However, this assumption does not mean that objects can be completely reconstructed.

1.1. Related Work

Existing shape completion approaches make use of geometric information represented at either low-level or high-level. Low-level geometry describes local structures, e.g. local smoothness, and can be used to fill small holes on broken surfaces. For example, Curless and Levoy [5] proposed to extract surfaces by examining the boundary of unseen and empty voxels. However, this method requires additional range images to carve away redundant surfaces. In [7], Davis et al. filled gaps and holes on broken surfaces by performing a convolution on the signed distance values. This process was repeated until a new implicit surface could be defined at the gaps. In [16], a broken object was represented in an octree grid on which inner and outer grid points were determined. The broken object was then constructed by contouring the grid points. In [29], holes on a broken

*This work was conducted when Duc Thanh Nguyen was working at the Singapore University of Technology and Design.

Density threshold: 20

A Field Model for Repairing 3D Shapes*

Duc Thanh Nguyen^{1,2}, Binh-Son Hua², Minh-Khoi Tran², Quang-Hieu Pham², and Sai-Kit Yeung²

¹School of Information Technology, Deakin University, Australia

²Singapore University of Technology and Design, Singapore

Abstract

This paper proposes a field model for repairing 3D shapes constructed from multi-view RGB data. Specifically, we represent a 3D shape in a Markov random field (MRF) in which the geometric information is encoded by random binary variables and the appearance information is retrieved from a set of RGB images captured at multiple viewpoints. The local priors in the MRF model capture the local structures of object shapes and are learnt from 3D shape templates using a convolutional deep belief network. Repairing a 3D shape is formulated as the maximum a posteriori (MAP) estimation in the corresponding MRF. Variational mean field approximation technique is adopted for the MAP estimation. The proposed method was evaluated on both artificial data and real data obtained from reconstruction of practical scenes. Experimental results have shown the robustness and efficiency of the proposed method in repairing noisy and incomplete 3D shapes.

1. Introduction

Suppose we are given a set of RGB/RGB-D images of an object captured at multiple viewpoints. The object in the real world (i.e. 3D space) is then re-constructed using some 3D reconstruction algorithm. Ideally, if an object can be observed in RGB/RGB-D images, it can be well reconstructed. However, in reality we have found that the reconstruction often fails even if the RGB/RGB-D data is complete. This is because the matching of the RGB data in structure-from-motion based reconstruction methods (e.g. [14]) could not be done accurately, specially for objects of uniform colours. For reconstruction methods using depth (e.g. [39, 4]), the missing of depth could also cause the incompleteness. We illustrate several cases of this situation in Fig. 1.

Recent advances of 3D acquisition devices and 3D scene reconstruction research [28, 38, 39, 40, 4] have enabled large-scale acquisition of 3D scene data and this has raised

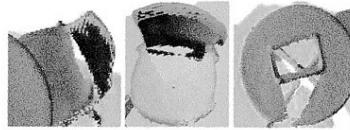


Figure 1. Examples of incomplete shapes after reconstruction using [4].

a demand on 3D data analysis. However, it often happens that the 3D data cannot be obtained at high quality (as shown in Fig. 1), even by recent reconstruction methods, e.g. [4]. Specifically, the 3D surfaces are missing and/or broken and this phenomenon causes difficulties for many sequential tasks such as 3D object detection and recognition [30, 36], shape analysis [20, 19], and scene understanding [32, 12]. Repairing missing and broken surfaces thus plays a critical role and deserves in-depth study. In this paper, we focus on repairing incomplete 3D shapes. This problem can be also referred to as shape completion. We assume objects are not occluded, i.e. they can be fully observed in RGB/RGB-D images. However, this assumption does not mean that objects can be completely reconstructed.

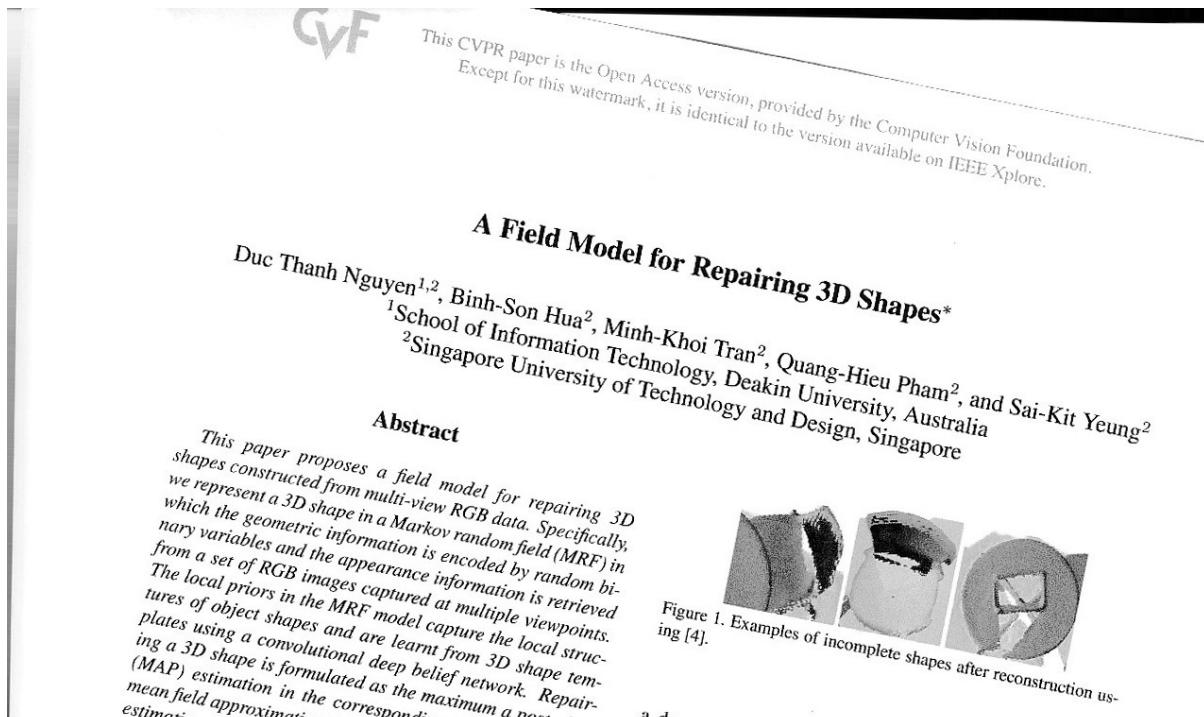
1.1. Related Work

Existing shape completion approaches make use of geometric information represented at either low-level or high-level. Low-level geometry describes local structures, e.g. local smoothness, and can be used to fill small holes on broken surfaces. For example, Curless and Levoy [5] proposed to extract surfaces by examining the boundary of unseen and empty voxels. However, this method requires additional range images to carve away redundant surfaces. In [7], Davis et al. filled gaps and holes on broken surfaces by performing a convolution on the signed distance values. This process was repeated until a new implicit surface could be defined at the gaps. In [16], a broken object was represented in an octree grid on which inner and outer grid points were determined. The broken object was then constructed by contouring the grid points. In [29], holes on a broken

*This work was conducted when Duc Thanh Nguyen was working at the Singapore University of Technology and Design.

Strategy: C) Point with maximum y-coordinate in each connected component.

Density Threshold: 5



1. Introduction

Suppose we are given a set of RGB/RGB-D images of an object captured at multiple viewpoints. The object in the real world (i.e. 3D space) is then re-constructed using some 3D reconstruction algorithm. Ideally, if an object can be observed in RGB/RGB-D images, it can be well reconstructed. However, in reality we have found that the reconstruction often fails even if the RGB/RGB-D data is complete. This is because the matching of the RGB data in structure-from-motion based reconstruction methods (e.g. [14]) could not be done accurately, specially for objects of uniform colours. For reconstruction methods using depth (e.g. [39, 4]), the missing of depth could also cause the incompleteness. We illustrate several cases of this situation in Fig. 1.

Recent advances of 3D acquisition devices and 3D scene reconstruction research [28, 38, 39, 40, 4] have enabled large-scale acquisition of 3D scene data and this has raised

Figure 1. Examples of incomplete shapes after reconstruction using [4].

a demand on 3D data analysis. However, it often happens that the 3D data cannot be obtained at high quality (as shown in Fig. 1), even by recent reconstruction methods, e.g. [4]. Specifically, the 3D surfaces are missing and/or broken and this phenomenon causes difficulties for many sequential tasks such as 3D object detection and recognition [30, 36], shape analysis [20, 19], and scene understanding [32, 12]. Repairing missing and broken surfaces thus plays a critical role and deserves in-depth study. In this paper, we focus on repairing incomplete 3D shapes. This problem can be also referred to as shape completion. We assume objects are not occluded, i.e. they can be fully observed in RGB/RGB-D images. However, this assumption does not mean that objects can be completely reconstructed.

1.1. Related Work

Existing shape completion approaches make use of geometric information represented at either low-level or high-level. Low-level geometry describes local structures, e.g. local smoothness, and can be used to fill small holes on broken surfaces. For example, Curless and Levoy [5] proposed to extract surfaces by examining the boundary of unseen and empty voxels. However, this method requires additional range images to carve away redundant surfaces. In [7], Davis et al. filled gaps and holes on broken surfaces by performing a convolution on the signed distance values. This process was repeated until a new implicit surface could be defined at the gaps. In [16], a broken object was represented in an octree grid on which inner and outer grid points were determined. The broken object was then constructed by contouring the grid points. In [29], holes on a broken

*This work was conducted when Duc Thanh Nguyen was working at the Singapore University of Technology and Design.

Density threshold: 10

A Field Model for Repairing 3D Shapes*

Duc Thanh Nguyen^{1,2}, Binh-Son Hoi¹
¹School of Information Technology, Deakin University
²Singapore University of Technology and Design
Abstract
 This paper proposes a field model for repairing 3D multi-view RGB data. Specifically, a Markov random field (MRF) based on multi-view RGB data is retrieved from a network. The MRF is encoded by random fields. The information is retrieved at multiple local viewpoints. The information is used as multiple local viewpoints to repair the damage in the 3D shape. The damage is captured from 3D shape.

Abstract This paper proposes a field model for repairing 3D shapes constructed from multi-view RGB data. Specifically, we represent the geometric information encoded by random binary variables and the appearance information is retrieved from a set of RGB images in the MRF model capture a 3D shape template of object shapes in the corresponding MRF. Variational priors using a convolutional network are adopted for the MAP estimation in the reconstruction technique was evaluated on both mean field approximation and the proposed method in repairing a 3D shape is formulated as the maximum a posteriori (MAP) estimation. The proposed method obtained results have shown the estimation of practical scenes. Experimental results have shown the robustness and efficiency of the proposed method in repairing noisy and incomplete 3D shapes.

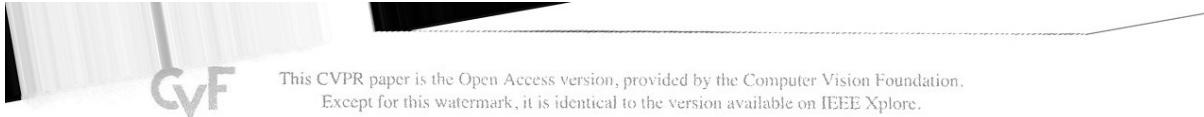
reconstruction have shown the need in repair-
ment on both local and global levels. For example, faces by themselves do not contain enough information to reconstruct them. This is because faces are often occluded by other objects or parts of the scene. In such cases, it is necessary to use additional information such as depth maps (RGB-D images) to reconstruct the faces accurately.

1.1. Related Work

Related Work Existing shape completion approaches make use of geometric information geometry described at either local structures, e.g. metric low-level geometry and can be used to fill small holes on broken surfaces. For example, Curless and Levoy [5] proposed to extract voxels to carve away local smoothness and can be used to fill gaps and holes on broken surfaces. However, this method requires additional range images to represent implicit surfaces. In [16], a broken object was represented by performing a convolution until a new inner and outer grid points were determined. The grid points were determined by contouring the grid points. This process was repeated until a octree grid on which object was then constructed. In [29], holes on a broken object were filled by examining the signed distance values on broken surfaces. However, this method requires additional range images to represent implicit surfaces. In [16], a broken object was represented by performing a convolution until a new inner and outer grid points were determined. The grid points were determined by contouring the grid points. This process was repeated until a octree grid on which object was then constructed.

and this has
and Design.

Density Threshold: 15



A Field Model for Repairing 3D Shapes*

Duc Thanh Nguyen^{1,2}, Binh-Son Hua², Minh-Khoi Tran², Quang-Hieu Pham², and Sai-Kit Yeung²

¹School of Information Technology, Deakin University, Australia

²Singapore University of Technology and Design, Singapore

Abstract

This paper proposes a field model for repairing 3D shapes constructed from multi-view RGB data. Specifically, we represent a 3D shape in a Markov random field (MRF) in which the geometric information is encoded by random binary variables and the appearance information is retrieved from a set of RGB images captured at multiple viewpoints. The local priors in the MRF model capture the local structures of object shapes and are learnt from 3D shape templates using a convolutional deep belief network. Repairing a 3D shape is formulated as the maximum a posteriori (MAP) estimation in the corresponding MRF. Variational mean-field approximation technique is adopted for the MAP estimation. The proposed method was evaluated on both artificial data and real data obtained from reconstruction of practical scenes. Experimental results have shown the robustness and efficiency of the proposed method in repairing noisy and incomplete 3D shapes.

1. Introduction

Suppose we are given a set of RGB/RGB-D images of an object captured at multiple viewpoints. The object in the real world (i.e. 3D space) is then re-constructed using some 3D reconstruction algorithm. Ideally, if an object can be observed in RGB/RGB-D images, it can be well reconstructed. However, in reality we have found that the reconstruction often fails even if the RGB/RGB-D data is complete. This is because the matching of the RGB data in structure-from-motion based reconstruction methods (e.g. [14]) could not be done accurately, specially for objects of uniform colours. For reconstruction methods using depth (e.g. [39, 4]), the missing of depth could also cause the incompleteness. We illustrate several cases of this situation in Fig. 1.

Recent advances of 3D acquisition devices and 3D scene reconstruction research [28, 38, 39, 40, 4] have enabled large-scale acquisition of 3D scene data and this has raised

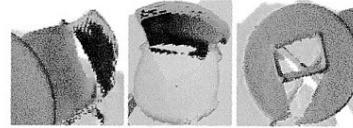


Figure 1. Examples of incomplete shapes after reconstruction using [4].

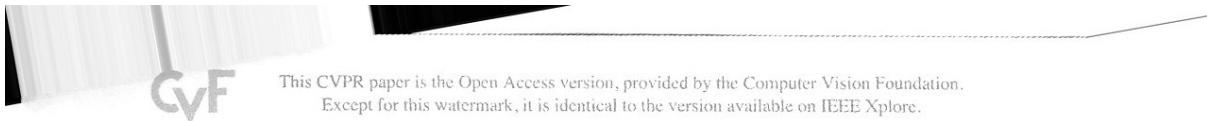
a demand on 3D data analysis. However, it often happens that the 3D data cannot be obtained at high quality (as shown in Fig. 1), even by recent reconstruction methods, e.g. [4]. Specifically, the 3D surfaces are missing and/or broken and this phenomenon causes difficulties for many sequential tasks such as 3D object detection and recognition [30, 36], shape analysis [20, 19], and scene understanding [32, 12]. Repairing missing and broken surfaces thus plays a critical role and deserves in-depth study. In this paper, we focus on repairing incomplete 3D shapes. This problem can be also referred to as shape completion. We assume objects are not occluded, i.e. they can be fully observed in RGB/RGB-D images. However, this assumption does not mean that objects can be completely reconstructed.

1.1. Related Work

Existing shape completion approaches make use of geometric information represented at either low-level or high-level. Low-level geometry describes local structures, e.g. local smoothness, and can be used to fill small holes on broken surfaces. For example, Curless and Levoy [5] proposed to extract surfaces by examining the boundary of unseen and empty voxels. However, this method requires additional range images to carve away redundant surfaces. In [7], Davis et al. filled gaps and holes on broken surfaces by performing a convolution on the signed distance values. This process was repeated until a new implicit surface could be defined at the gaps. In [16], a broken object was represented in an octree grid on which inner and outer grid points were determined. The broken object was then constructed by contouring the grid points. In [29], holes on a broken

*This work was conducted when Duc Thanh Nguyen was working at the Singapore University of Technology and Design.

Density threshold: 20



A Field Model for Repairing 3D Shapes*

Duc Thanh Nguyen^{1,2}, Binh-Son Hua², Minh-Khoi Tran², Quang-Hieu Pham², and Sai-Kit Yeung²

¹School of Information Technology, Deakin University, Australia

²Singapore University of Technology and Design, Singapore

Abstract

This paper proposes a field model for repairing 3D shapes constructed from multi-view RGB data. Specifically, we represent a 3D shape in a Markov random field (MRF) in which the geometric information is encoded by random binary variables and the appearance information is retrieved from a set of RGB images captured at multiple viewpoints. The local priors in the MRF model capture the local structures of object shapes and are learnt from 3D shape templates using a convolutional deep belief network. Repairing a 3D shape is formulated as the maximum a posteriori (MAP) estimation in the corresponding MRF. Variational mean field approximation technique is adopted for the MAP estimation. The proposed method was evaluated on both artificial data and real data obtained from reconstruction of practical scenes. Experimental results have shown the robustness and efficiency of the proposed method in repairing noisy and incomplete 3D shapes.

1. Introduction

Suppose we are given a set of RGB/RGB-D images of an object captured at multiple viewpoints. The object in the real world (i.e. 3D space) is then re-constructed using some 3D reconstruction algorithm. Ideally, if an object can be observed in RGB/RGB-D images, it can be well reconstructed. However, in reality we have found that the reconstruction often fails even if the RGB/RGB-D data is complete. This is because the matching of the RGB data in structure-from-motion based reconstruction methods (e.g. [14]) could not be done accurately, specially for objects of uniform colours. For reconstruction methods using depth (e.g. [39, 4]), the missing of depth could also cause the incompleteness. We illustrate several cases of this situation in Fig. 1.

Recent advances of 3D acquisition devices and 3D scene reconstruction research [28, 38, 39, 40, 4] have enabled large-scale acquisition of 3D scene data and this has raised

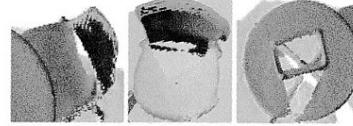


Figure 1. Examples of incomplete shapes after reconstruction using [4].

a demand on 3D data analysis. However, it often happens that the 3D data cannot be obtained at high quality (as shown in Fig. 1), even by recent reconstruction methods, e.g. [4]. Specifically, the 3D surfaces are missing and/or broken and this phenomenon causes difficulties for many sequential tasks such as 3D object detection and recognition [30, 36], shape analysis [20, 19], and scene understanding [32, 12]. Repairing missing and broken surfaces thus plays a critical role and deserves in-depth study. In this paper, we focus on repairing incomplete 3D shapes. This problem can be also referred to as shape completion. We assume objects are not occluded, i.e. they can be fully observed in RGB/RGB-D images. However, this assumption does not mean that objects can be completely reconstructed.

1.1. Related Work

Existing shape completion approaches make use of geometric information represented at either low-level or high-level. Low-level geometry describes local structures, e.g. local smoothness, and can be used to fill small holes on broken surfaces. For example, Curless and Levoy [5] proposed to extract surfaces by examining the boundary of unseen and empty voxels. However, this method requires additional range images to carve away redundant surfaces. In [7], Davis et al. filled gaps and holes on broken surfaces by performing a convolution on the signed distance values. This process was repeated until a new implicit surface could be defined at the gaps. In [16], a broken object was represented in an octree grid on which inner and outer grid points were determined. The broken object was then constructed by contouring the grid points. In [29], holes on a broken

*This work was conducted when Duc Thanh Nguyen was working at the Singapore University of Technology and Design.

Speed of Computational Experiments Conducted:

2.1 Candidate Point Selection Analysis:

Strategy a:

Estimated angle: 6.00 degrees
Selection time: 0.0090 seconds
Hough transform time: 0.0295 seconds
Total time: 0.1579 seconds

Strategy b:

Estimated angle: 15.00 degrees
Selection time: 0.0000 seconds
Hough transform time: 0.0024 seconds
Total time: 0.0071 seconds

Strategy c:

Estimated angle: 31.00 degrees
Selection time: 18.1280 seconds
Hough transform time: 0.0024 seconds
Total time: 18.1348 seconds

2.2 Parameter Setting Analysis:

Density Threshold: 5

Strategy a:

Estimated angle: 5.00 degrees
Selection time: 0.0060 seconds
Hough transform time: 0.0302 seconds
Total time: 0.1421 seconds

Strategy b:

Estimated angle: 1.00 degrees
Selection time: 0.0000 seconds
Hough transform time: 0.0032 seconds
Total time: 0.0098 seconds

Strategy c:

Estimated angle: 1.00 degrees
Selection time: 17.9041 seconds
Hough transform time: 0.0033 seconds
Total time: 17.9142 seconds

Density Threshold: 10

Strategy a:

Estimated angle: 6.00 degrees
Selection time: 0.0057 seconds
Hough transform time: 0.0308 seconds
Total time: 0.1463 seconds

Strategy b:

Estimated angle: 15.00 degrees
Selection time: 0.0000 seconds
Hough transform time: 0.0022 seconds
Total time: 0.0066 seconds

Strategy c:

Estimated angle: 31.00 degrees
Selection time: 17.9144 seconds
Hough transform time: 0.0023 seconds
Total time: 17.9212 seconds

Density Threshold: 15

Strategy a:

Estimated angle: 6.00 degrees
Selection time: 0.0059 seconds
Hough transform time: 0.0306 seconds
Total time: 0.1440 seconds

Strategy b:

Estimated angle: -10.00 degrees
Selection time: 0.0000 seconds
Hough transform time: 0.0022 seconds
Total time: 0.0068 seconds

Strategy c:

Estimated angle: -10.00 degrees
Selection time: 17.8879 seconds
Hough transform time: 0.0023 seconds
Total time: 17.8945 seconds

Density Threshold: 20

Strategy a:

Estimated angle: 7.00 degrees
Selection time: 0.0057 seconds
Hough transform time: 0.0294 seconds
Total time: 0.1378 seconds

Strategy b:

Estimated angle: -10.00 degrees
Selection time: 0.0000 seconds
Hough transform time: 0.0022 seconds
Total time: 0.0062 seconds

Strategy c:

Estimated angle: -10.00 degrees
Selection time: 17.8826 seconds
Hough transform time: 0.0024 seconds
Total time: 17.8891 seconds

The analysis reveals variations in performance and accuracy across different candidate point selection strategies and density thresholds. Visual inspection of the deskewed images provides crucial information that supports and extends the numerical analysis.

- **Strategy 'a' (All foreground pixels):**

- Consistently estimates the skew angle between 5-7 degrees
 - Moderate processing times (0.14-0.17 seconds)
 - Visual inspection confirms this strategy produces accurate deskewing results
- **Strategy 'b' (All connected components centres):**
 - Extremely fast (0.006-0.01 seconds)
 - Inconsistent angle estimations ranging from 1 to 15 degrees
 - Visual inspection shows good deskewing results, despite the inconsistency in angle estimation
- **Strategy 'c' (Point with maximum y-coordinate in each connected component):**
 - Slowest by far (17.9-18.8 seconds)
 - Widely varying angle estimates (1 to 31 degrees)
 - Visual inspection shows good deskewing results, despite the inconsistency in angle estimation and long processing time

Density threshold adjustments appear to have a minimal impact on processing times but can affect the estimated angles, particularly for strategies 'b' and 'c'. The inconsistency in angle estimations across different strategies and thresholds confirms that the skew correction method's reliability depends on the specific characteristics of the input document.

Visual inspection of the deskewed images shows that:

- All strategies produce well-aligned documents that are easily readable.
- Despite the numerical inconsistencies in angle estimation for strategies 'b' and 'c', their visual results are comparable to strategy 'a' for this document.

Chosen Strategy:

Based on both the numerical results and visual inspection, strategy "a" (using all foreground pixels as candidate points) remains the optimal choice. The density threshold of 20, as originally determined, appears to provide the best balance between accuracy and processing time. The deskewed image using the strategy 'a' with a threshold of 20 shows proper alignment and readability.

This choice is supported by:

- Consistent angle estimation for strategy 'a'
- Acceptable processing time
- Visual confirmation of well-aligned, readable results
- The original analysis indicated that a threshold of 20 yielded the best outcome

While strategies 'b' and 'c' can produce visually similar results in this case, their inconsistencies in processing time and angle estimation make them less reliable options for a wider range of documents. However, if processing speed is a critical factor, strategy 'b' might be considered as an alternative, with the caveat that its performance may vary more across different documents.

3. Test images and Results:



Image 1: Skew = 46.00° , Selection Time = 0.0575s, Points = 4185914

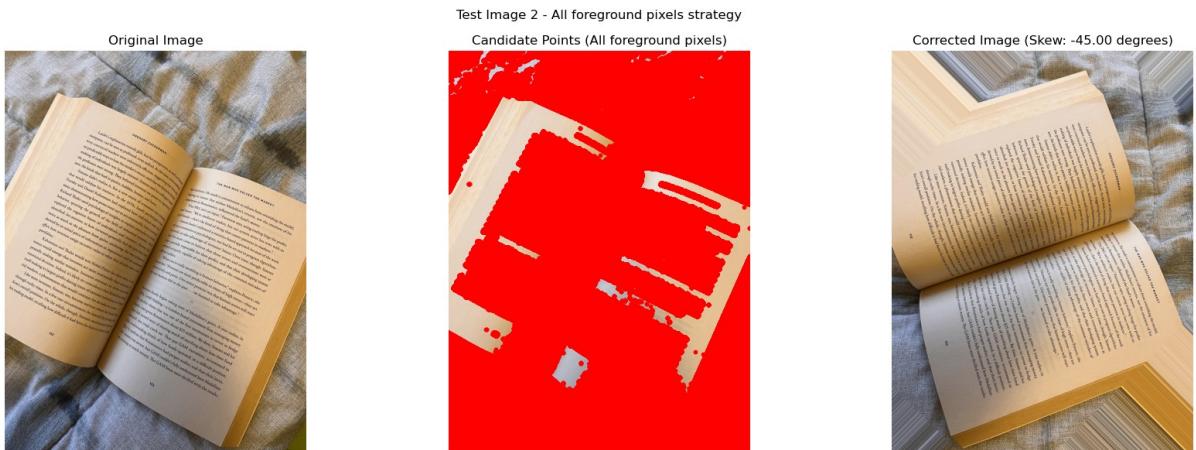


Image 2: Skew = -45.00° , Selection Time = 0.0684s, Points = 5653923



Image 3: Skew = 46.00° , Selection Time = 0.0569s, Points = 4128387

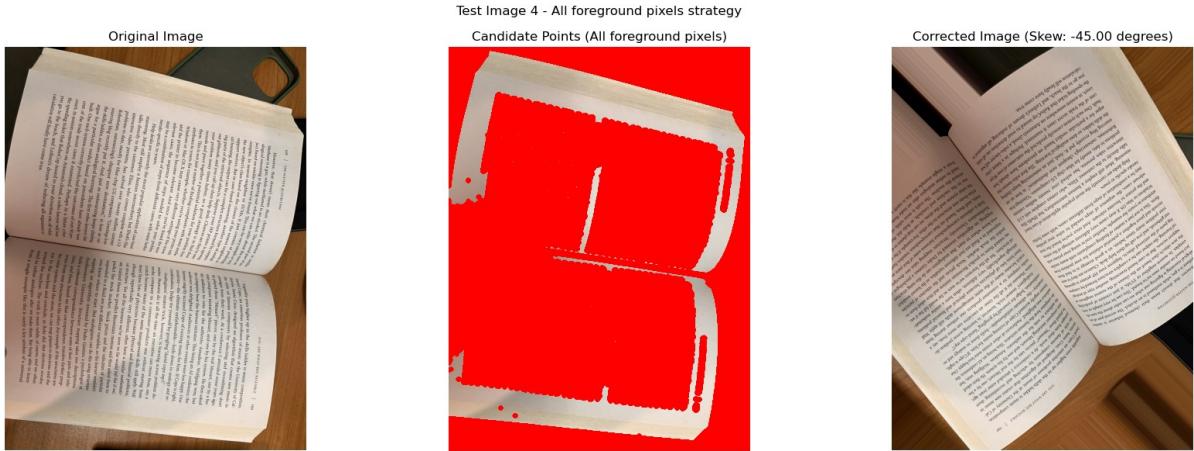


Image 4: Skew = -45.00° , Selection Time = 0.0468s, Points = 4275403

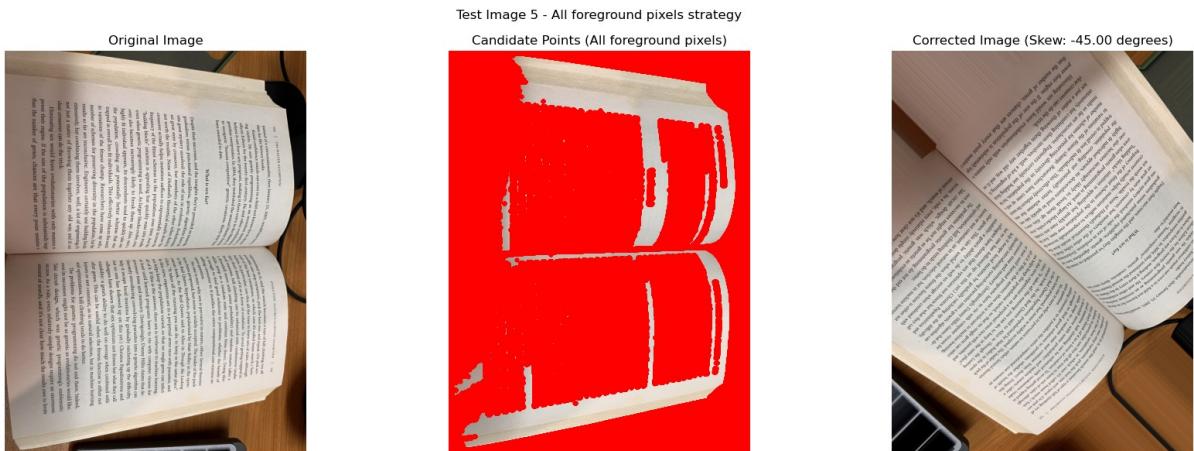


Image 5: Skew = -45.00° , Selection Time = 0.0586s, Points = 4628999

The algorithm successfully detected significant skew angles of either 45 or 46 degrees (positive or negative) for all five images, indicating its effectiveness in identifying large skew angles. Processing times for candidate point selection were relatively uniform, ranging from 0.0407 to 0.0606 seconds, suggesting efficient implementation. The number of candidate points varied between 3.7 to 4.6 million across images, reflecting the all-foreground-pixels strategy. Visual inspection of the corrected images confirms the algorithm's ability to effectively straighten severely skewed documents, demonstrating its robustness in handling various book and document orientations.

4. Recognized text from doc.jpg:

```
1 Recognized Text:  
2 This CVPE  
3 |  
4 A Field Model for Repairing 3D Shapes*  
5  
6 Duc Thanh Nguyen', Binh-Son Hua, Minh-Khoi Tran", Quang-Hieu Pham', and Sai-Kit Yeung?  
7 1School of Information Technology, Deakin University, Australia  
8 ?Singapore University of Technology and Design, Singapore  
9  
10 Abstract  
11  
12 This paper proposes a field model for repairing 3D  
13 shapes constructed from multi-view RGB data. Specifically,  
14 we represent a 3D shape in a Markov random field (MRF) in  
15 which the geometric information is encoded by random bi-  
16 nary variables and the appearance information is retrieved  
17 from a set of RGB images captured at multiple viewpoints.  
18 The local priors in the MRF model capture the local struc-  
19 tures of object shapes and are learnt from 3D shape tem-  
20 plates using a convolutional deep belief network. Repair-  
21 ing a 3D shape is formulated as the maximum a posteriori  
22 (MAP) estimation in the corresponding MRF. Variational  
23 mean field approximation technique is adopted for the MAP  
24 estimation. The proposed method was evaluated on both  
25 artificial data and real data obtained from reconstruction  
26 of practical scenes. Experimental results have shown the  
27 robustness and efficiency of the proposed method in repair-  
28 ing noisy and incomplete 3D shapes.  
29  
30 1. Introduction  
31  
32 Suppose we are given a set of RGB/RGB-D images of  
33 an object captured at multiple viewpoints. The object in the  
34 real world (i.e. 3D space) is then re-constructed using some  
35 3D reconstruction algorithm. Ideally, if an object can be ob-  
36 served in RGB/RGB-D images, it can be well reconstructed.  
37 However, in reality we have found that the reconstruction  
38 often fails even if the RGB/RGB-D data is complete. This  
39 is because the matching of the RGB data in structure-from-  
40 motion based reconstruction methods (e.g. [14]) could not  
41 be done accurately, specially for objects of uniform colours.  
42 For reconstruction methods using depth (e.g. [39, 4]), the  
43 missing of depth could also cause the incompleteness. We  
44 illustrate several cases of this situation in Fig. 1.  
45  
46 Recent advances of 3D acquisition devices and 3D scene  
47 reconstruction research [28, 38, 39, 40, 4] have enabled
```

Recognized text from deskewed doc.jpg:

```
91 Recognized Text:  
92 A Field Model for Repairing 3D Shapes*  
93  
94 Duc Thanh Nguyen', Binh-Son Hua?, Minh-Khoi Tran", Quang-Hieu Pham?, and Sai-Kit Yeung?  
95  
96 1$chool of Information Technology, Deakin University, Australia  
97 ?Singapore University of Technology and Design, Singapore  
98  
99 Abstract  
100  
101 This paper proposes a field model for repairing 3D  
102 shapes constructed from multi-view RGB data. Specifically,  
103 we represent a 3D shape in a Markov random field (MRF) in  
104 which the geometric information is encoded by random bi-  
105 nary variables and the appearance information is retrieved  
106 from a set of RGB images captured at multiple viewpoints.  
107 The local priors in the MRF model capture the local struc-  
108 tures of object shapes and are learnt from 3D shape tem-  
109 plates using a convolutional deep belief network. Repair-  
110 ing a 3D shape is formulated as the maximum a posteriori  
111 (MAP) estimation in the corresponding MRF. Variational  
112 mean field approximation technique is adopted for the MAP  
113 estimation. The proposed method was evaluated on both  
114 artificial data and real data obtained from reconstruction  
115 of practical scenes. Experimental results have shown the  
116 robustness and efficiency of the proposed method in repair-  
117 ing noisy and incomplete 3D shapes.  
118  
119 1. Introduction  
120  
121 Suppose we are given a set of RGB/RGB-D images of  
122 an object captured at multiple viewpoints. The object in the  
123 real world (i.e. 3D space) is then re-constructed using some  
124 3D reconstruction algorithm. Ideally, if an object can be ob-  
125 served in RGB/RGB-D images, it can be well reconstructed.  
126 However, in reality we have found that the reconstruction  
127 often fails even if the RGB/RGB-D data is complete. This  
128 is because the matching of the RGB data in structure-from-  
129 motion based reconstruction methods (e.g. [14]) could not  
130 be done accurately, specially for objects of uniform colours.  
131 For reconstruction methods using depth (e.g. [39, 4]), the  
132 missing of depth could also cause the incompleteness. We  
133 illustrate several cases of this situation in Fig. 1.  
134  
135 Recent advances of 3D acquisition devices and 3D scene  
136 reconstruction research [28, 38, 39, 40, 4] have enabled  
137 large-scale acquisition of 3D scene data and this has raised
```

For both the original and deskewed images, the system successfully extracted text content, focusing on the paper's title, authors, and abstract. The processing times were similar for both images (4.03 and 4.09 seconds respectively), indicating that the pre-deskewing didn't significantly impact overall performance. The recognized text appears largely accurate, capturing key details of the academic paper, though some minor errors are present (e.g., "This CVPE" at the start of the original image output). The consistent performance across both images suggests that the built-in skew correction is effective, as it produced similar results to the pre-deskewed image.