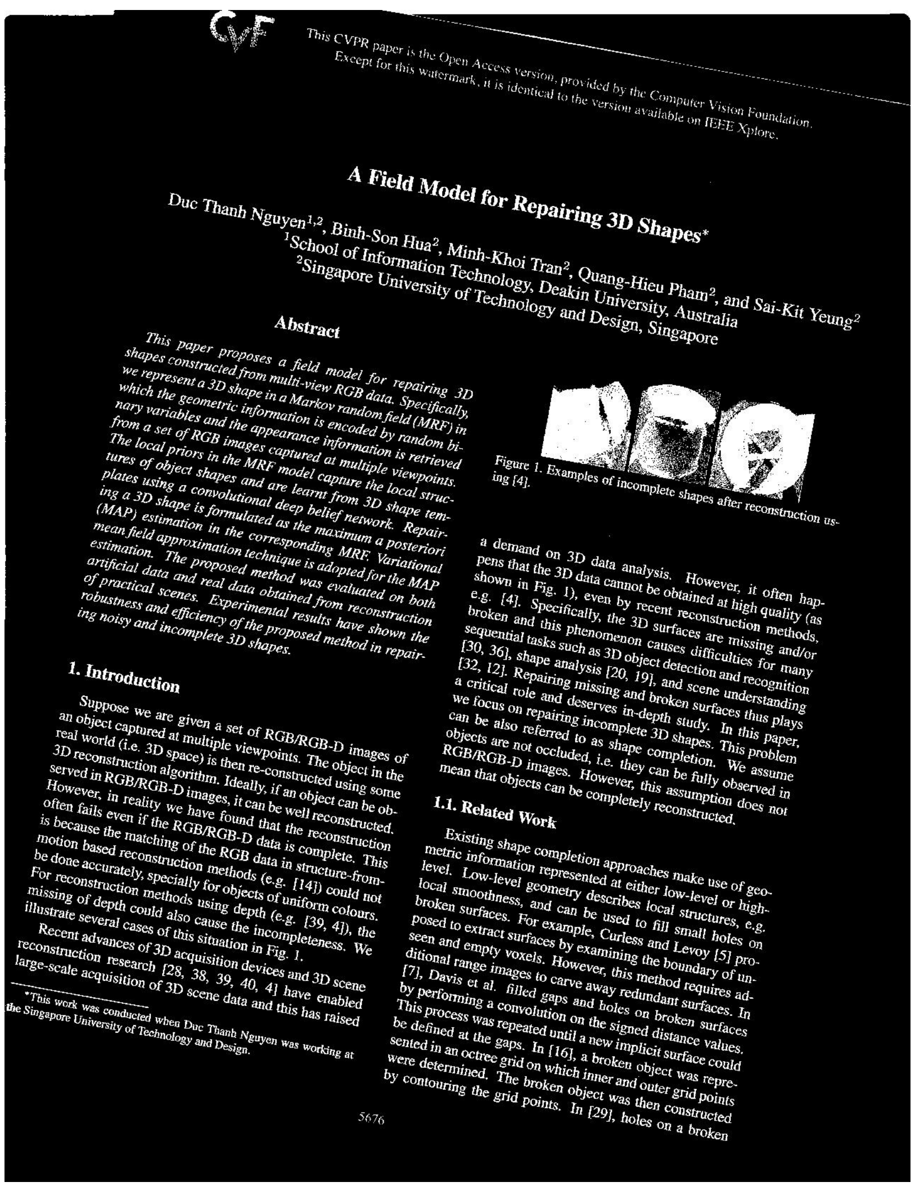
Distinction Task 2.3: Document analysis and recognition

1. For each candidate point selection strategy, write the negative image of doc.jpg with selected candidate points only (i.e., result of step 5 in Task 1) to image file.

Strategy 1



A Field Model for Repairing 3D Shapes*

Duc Thanh Nguyen[1,2], Binh-Son Hua[2], Minh-Khoi Tran[2], Quang-Hieu Pham[2], and Sai-Kit Yeung[2]
[1]School of Information Technology, Deakin University, Australia
[2]Singapore 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

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

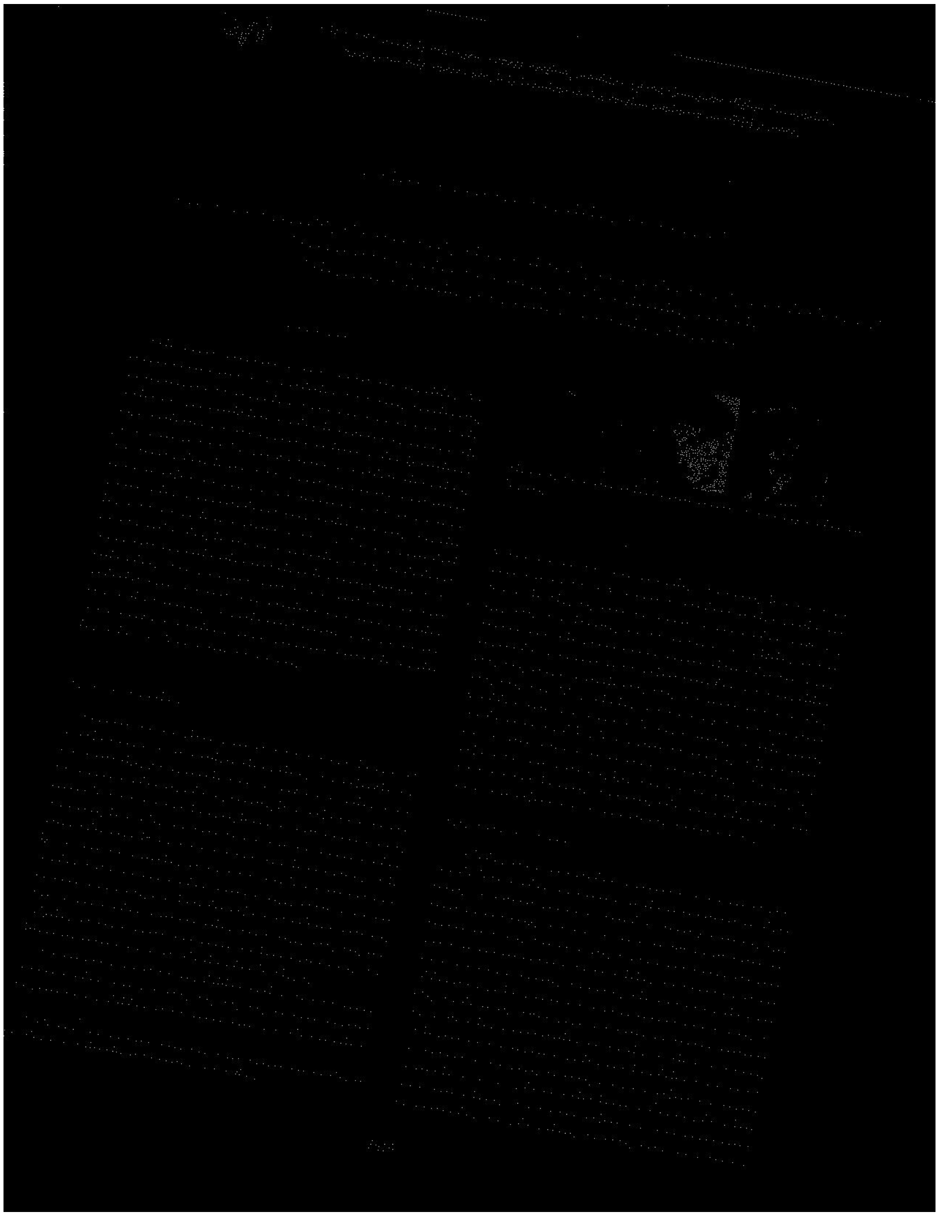
Figure 1. Examples of incomplete shapes after reconstruction using [4].

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

Strategy 2



Strategy 3

2. Write the deskewing results of doc.jpg using different candidate point selection strategies into image files.

Strategy 1

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 Hua[2], Minh-Khoi Tran[2], Quang-Hieu Pham[2], and Sai-Kit Yeung[2]
[1]School of Information Technology, Deakin University, Australia
[2]Singapore 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

*This work was conducted when Duc Thanh Nguyen was working at the 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

5676

Strategy 2

A Field Model for Repairing 3D Shapes*

Duc Thanh Nguyen[1,2], Binh-Son Hua[2], Minh-Khoi Tran[2], Quang-Hieu Pham[2], and Sai-Kit Yeung[2]
[1]School of Information Technology, Deakin University, Australia
[2]Singapore 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

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

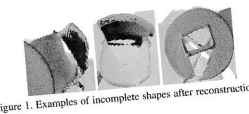


Figure 1. Examples of incomplete shapes after reconstruction using [4].

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

Strategy 3

A Field Model for Repairing 3D Shapes*

Duc Thanh Nguyen[1,2], Binh-Son Hua[2], Minh-Khoi Tran[2], Quang-Hieu Pham[2], and Sai-Kit Yeung[2]
[1]School of Information Technology, Deakin University, Australia
[2]Singapore 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.

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

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 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.

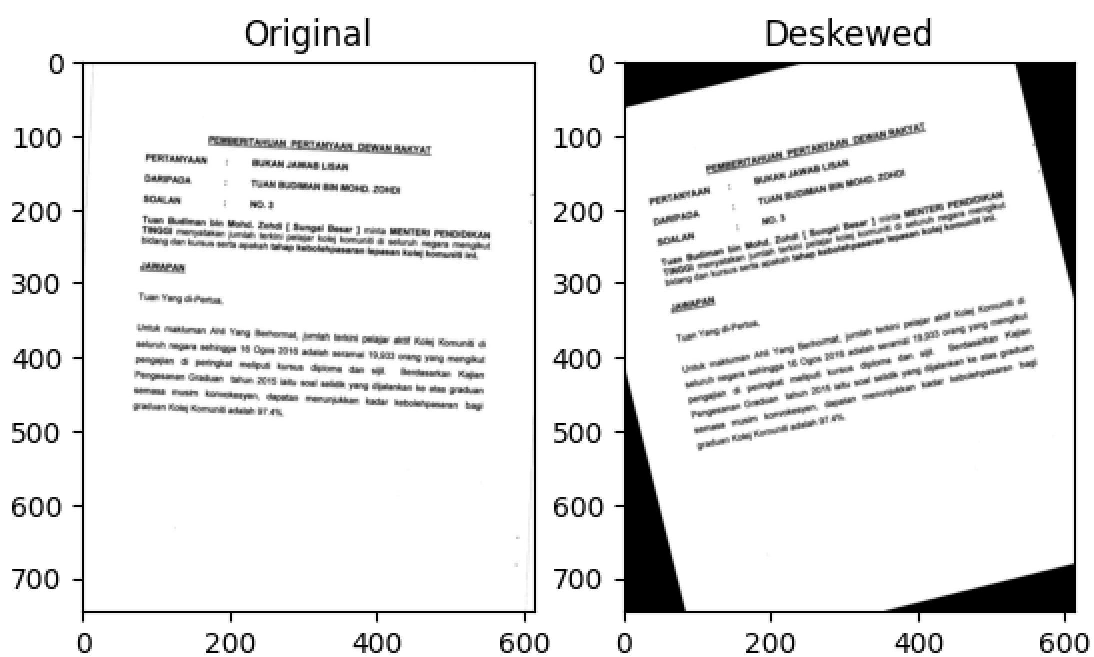
In the Hough Transform-based document skew estimation task, the computational complexity varies significantly across the three candidate point selection strategies.

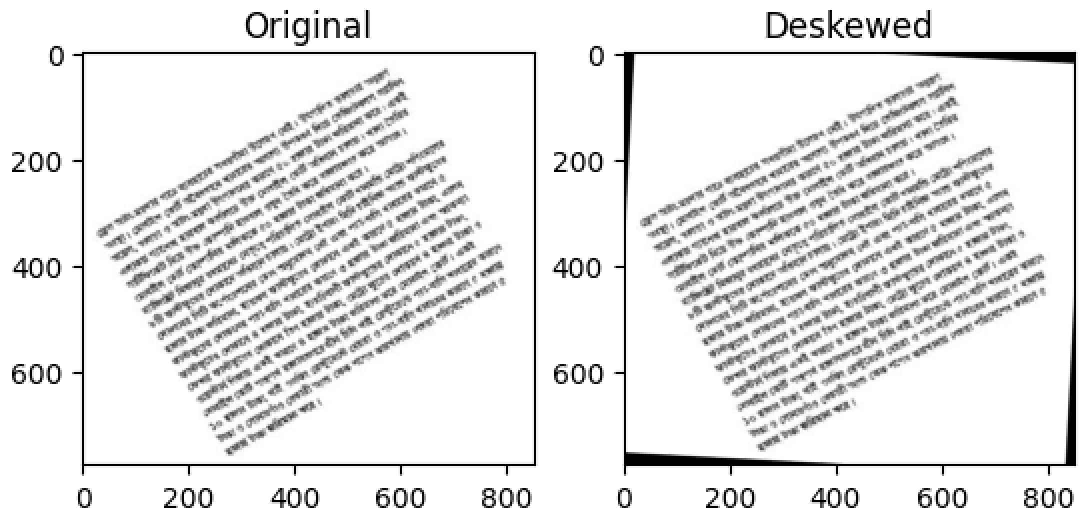
- **Strategy 1 (all foreground pixels): This has a complexity of O(i×j), where i and j are the dimensions of the image, as it processes every pixel.**
- **Strategy 2 (centroids of connected components): The complexity is O(C), where C is the number of connected components, making it more efficient than Strategy 1.**
- **Strategy 3 (maximum y-coordinate in each component): This involves O(C×p)), where p is the average number of points per component, leading to significantly higher complexity and runtime.**

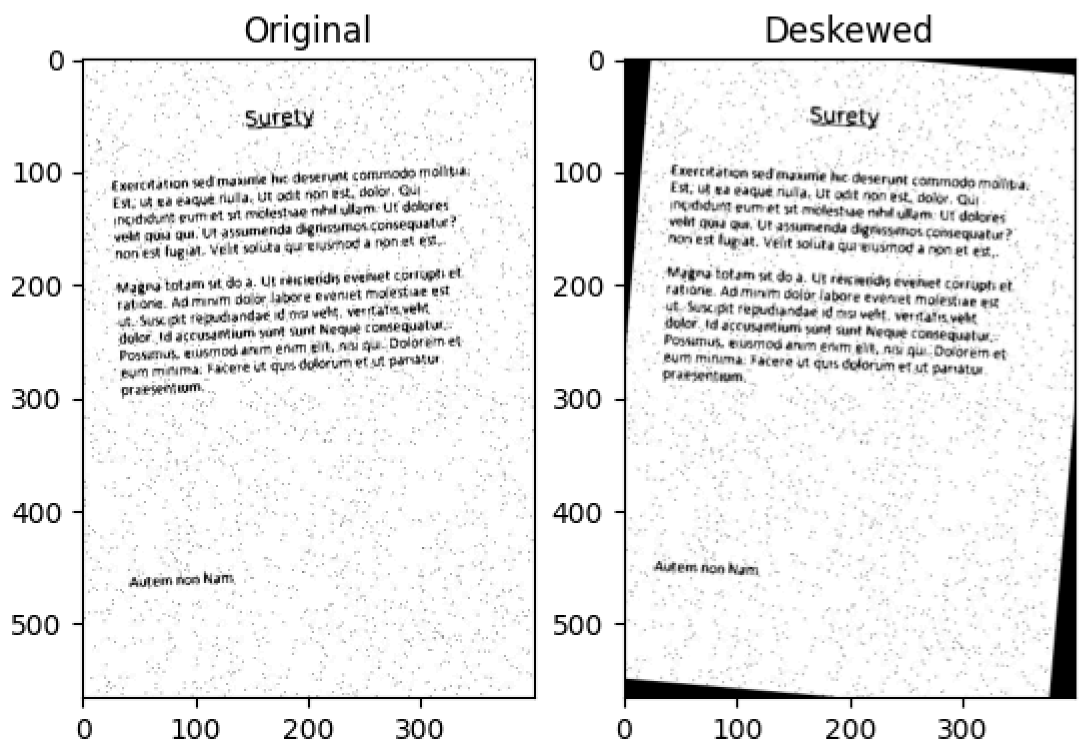
Time required for the strategies,Hough transform step (Step 7), and the entire process (including all the steps):

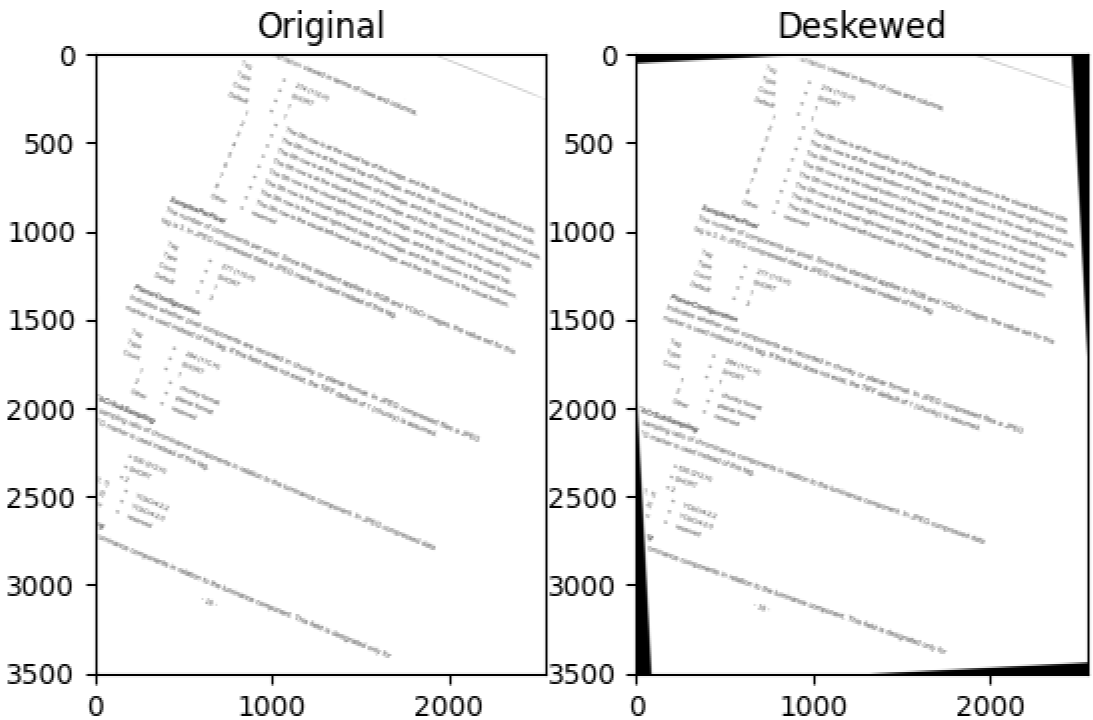
a.Time required for this strategy 1 is: 10.340595006942749 seconds
b.Time required for this strategy 2 is: 0.008671998977661133 seconds
c.Time required for this strategy 3 is: 35.51823973655701 seconds
d.Time required for this step 7 is: 0.9791727066040039 seconds
e.Time required for the whole process is: 38.66529679298401 seconds

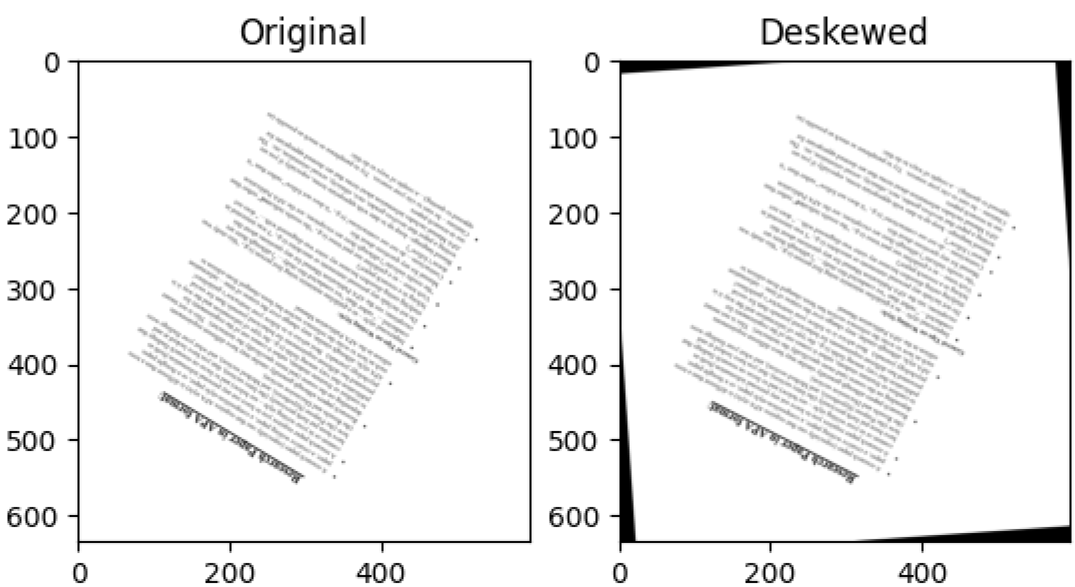
3. Write the deskewing results of images you have collected in Task 3 into image files.




Original
Deskewed
0
200
400
600
800


Original
Deskewed
Surety
Autem non Nam
0
100
200
300
400
500


Original
Deskewed
0
500
1000
1500
2000
2500
3000
3500
0
1000
2000


Original
Deskewed
0
100
200
300
400
500
600
0
200
400

4. Make screenshots of the recognised text (i.e., results of print(text)) for the skewed document image in doc.jpg and its deskewed version

Screenshots of the recognised text From Skewed Image

```
Recognized Text from Skewed Image:




A Fielg Mode fo Pairing 3D Shapes«
Uc Thanh Nguyen! #. Binh-g a? Mi Ol Tran?
1S 001 of 1, n

im", and §
eak: Niversit Ustralig
log id i





° 28S raised :
Sented in ai

ducted When Due Thanh Neuyen was Working ap

sity °F Technology and Design.
```

Screenshots of the recognised text From DeSkewed Image

```
Recognized Text from Deskewed Image:

A Field Model for Repairing 3D Shapes*

Duc Thanh Nguyen}, Binh-Son Hua?, Minh-Khoi Tran?, Quang-Hieu Pham®, and Sai-Kit Yeung?
"School of Information Technology, Deakin University, Australia
*Singapore 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 a3D shape ina Markov random field (MRF) in
which the Seometric information j. encoded by random bi-
"ary variables and the appearance information is retrieved
from a set of RGB images captured at multiple viewpoints,
The local priors in the MRE model capture the local struc.
tures of object Shapes and are learnt from 3D shape tem-
plates using a convolutional deep belief network. Repair-
ing a 3D shape is formulated as the maximum a posteriori
(MAP) estimation in the corresponding MRF. Variational
mean field approximation fechnique is adopted for the MAP
estimation. The Proposed method was evaluated on both
artificial data and real data obtained from reconstruction
of practical scenes, f.
robustness and efficiency of the proposed method in repair-
ing noisy and incomplete 3D shapes,
```

```
1. Introduction

Suppose we are Siven 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 ob-
Served in RGB/RGB-D images, it can be well reconstructed,
However, in reality we have found that the
often fails i



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.

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



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




Figure 1. Examples of incomplete shapes after reconstruction us-
ing [4]
```

```
S. However, it often hap-
Pens that the 3D data cannot be obtained at high quality (as
shown in Fig. 1), even by recent reconstruction methods,
ifically, the 31D Surfaces are missing and/or
broken and this Phenomenon Causes difficulties for many
Sequential tasks such as 3D object detection and recognition
[30, 36], s] ‘ape 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,







1.1. Related Work

Existing shape completion approaches make use of geo-
metric information Tepresented 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] pro-
Posed to extract surfaces by examining the boundary of un-
seen and empty voxels, However, this method requires ad-
ditional Tange 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 repre-
sented in an octree grid on which inner and outer grid points
Were determined. The broken object was then constructed
by contouring the &rid points. In [29], holes on @ broken

re
```

Analyses and discussions required in Tasks 2-4 to OnTrack

Does the Hough transform accurately work in every case? If not, what could be the reason and how to address it.

- In the process of dewskewing the image ,the density_threshold of 10 works the best for the doc.jpg to skew angle of the doc image, but for the threshold 50,it makes the type of the lines to None which creates problem in appending or collecting the angles thus the whole process is disturbed. So I will be choosing the threshold of 10 to deskewing the doc images.

There could be multiple reason for that and their solutions are described below:-

1. Noise:

- **Problem: Noise in the image can lead to false positives or disrupt the detection of shapes.**
- **Solution: Applying preprocessing techniques such as filtering (e.g., Gaussian blur) or edge detection (e.g., Canny edge detector) to reduce noise.**

2.Resolution and Scale:

- **Problem: The resolution of the image and the scale of the shapes can affect detection. Shapes that are too small or too large compared to the resolution may not be detected correctly.**

- **Solution: Ensuring that the image has an appropriate resolution and apply scaling techniques to match the expected size of the shapes.**

3. Proper Threshold Initialization:

- **Problem: Not Initializing proper threshold for determining the angles can also be a problem because the hough transform might detect wrong lines of the text with the other.**

A Field Model for Repairing 3D Shapes*

Duc Thanh Nguyen[1,2], Binh-Son Hua[2], Minh-Khoi Tran[2], Quang-Hieu Pham[2], and Sai-Kit Yeung[2]
[1]School of Information Technology, Deakin University, Australia
[2]Singapore 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 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.

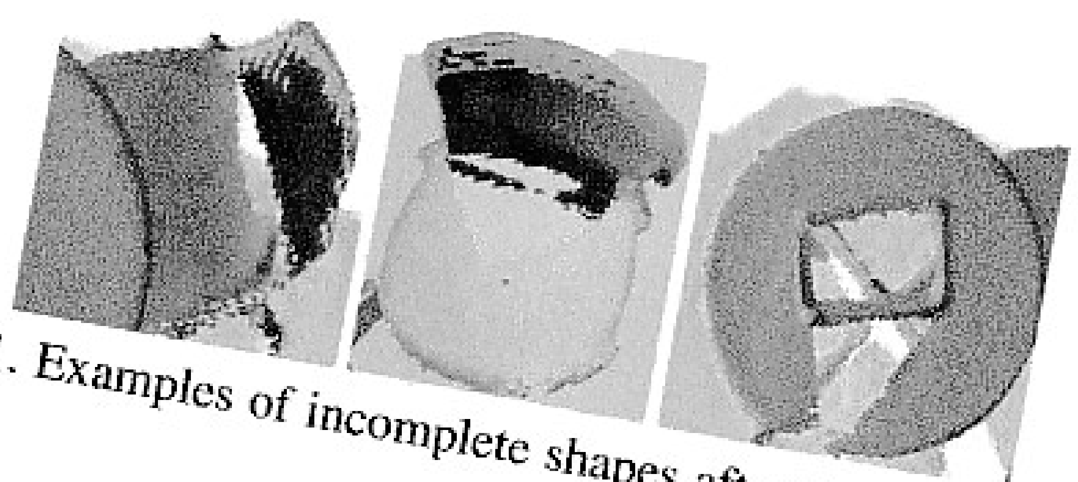


Figure 1. Examples of incomplete shapes after reconstruction using [4].

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.

A Field Model for Repairing 3D Shapes*

Duc Thanh Nguyen[1,2], Binh-Son Hua[2], Minh-Khoi Tran[2], Quang-Hieu Pham[2], and Sai-Kit Yeung[2]
[1]School of Information Technology, Deakin University, Australia
[2]Singapore 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 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.

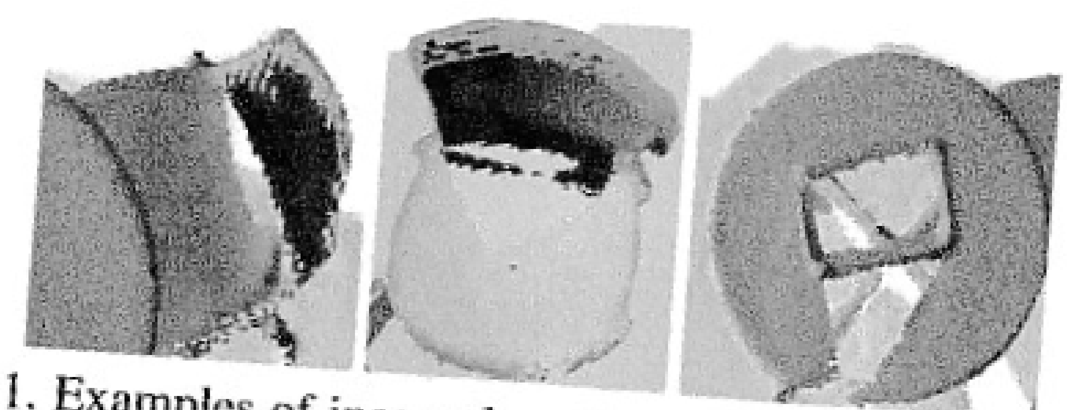


Figure 1. Examples of incomplete shapes after reconstruction using [4].

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.