

Buildings from Images
**Beispiel-basierte Rekonstruktion von Gebäuden aus einem
Einzelbild**

Thien Pham

Hochschule für Angewandte Wissenschaften Hamburg

1 Einführung und Motivation

Diese Ausarbeitung beschäftigt sich mit der Technik, aus zweidimensionalen Bildern dreidimensionale Objekte zu generieren. Hierbei werden speziell Gebäude wie Ein-, Mehrfamilienhäuser oder Hochhäuser in Betracht gezogen. Um aus einem Bild (Foto, Rendering) das dazugehörige 3D-Objekt rekonstruieren zu können, muss zunächst eine digitale Abbildung des Objekts (des Hauses) aus dem Bild extrahiert werden. Dafür gibt es in der Computergrafik verschiedene Chroma Key Algorithmen, mit denen sich unwichtige Elemente bzw. Farben in einem Bild ausschwärzen lassen (siehe Vergleich in Abbildung 1). Um die Kanten des Ergebnisses zu glätten, können zusätzlich Median Filter und morphologische Operatoren angewendet werden. Weiterhin können Ecken und Kanten des extrahierten Bildes per Bilderkennung gefunden werden, die dann für die Rekonstruktion des 3D-Objektes verwendet werden können.



Abb. 1. Links: Beispielbild eines Einfamilienhauses. Rechts: Das Einfamilienhaus wurde extrahiert und alles Unwichtige im Hintergrund ausgeschwärzt. (Quelle: [1])

Ähnliche Projekte wie diese finden sich bereits bei Apple (Apple Flyover, siehe Abbildung 2) und Google (Google Earth, siehe Abbildung 3). Die Objekterkennung in beiden Anwendungen basiert auf Fotografie. Die beiden Unternehmen ließen mit Kameras ausgestatteten Flugzeuge über verschiedene Städte fliegen. Dabei wurden während des Flugs kontinuierlich Fotos gemacht. Aus den entstandenen Fotos aus verschiedenen Perspektiven wurden die 3D-Gebäude rekonstruiert.

2 Zielstellung

Anders als bei Apple und Google ist das Ziel dieser Arbeit, aus einem oder maximal zwei Bildern das 3D-Objekt zu rekonstruieren. Dabei werden verschiedene Gebäude (Hochhäuser, Einfamilienhäuser, etc.) anhand von gegebenen zweidimensionalen Bildern (z.B. Fotos) zu 3D-Objekten zu rekonstruieren. Das heißt, es soll bis zum Abschluss der Masterarbeit ein Programm entwickelt werden, das ein Bild als Eingabe bekommt und die 3D-Grafik dieses Gebäudes erstellt. Das resultierende 3D-Objekt wird mithilfe einer Beispiel-basierten Rekonstruktion und einer Datenbank, die aus bereits vorhandenen 3D-Objekten besteht,

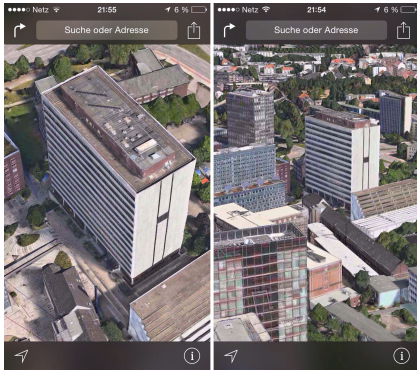


Abb. 2. Screenshot Apple Flyover
(Quelle: Apple Inc., Apple Maps)

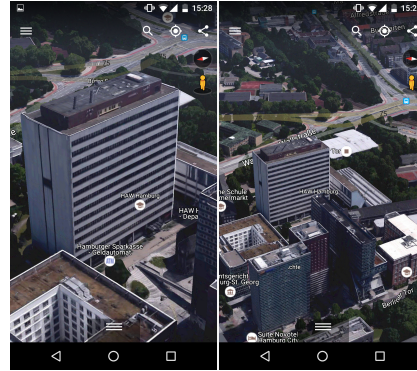


Abb. 3. Screenshot Google Earth
(Quelle: Google Inc., Google Earth)

erstellt. Bei dem Datenbank-Ansatz werden verschiedene Eigenschaften (Winkel, Rundungen, Abstände) des zu rekonstruierenden Objekts mit den Objekten aus der Datenbank verglichen.

3 Grundlagen: 3D-Objekte in der Computergrafik

Objekte in der Computergrafik bestehen aus folgenden Grundelementen: Punkte und Linien, die miteinander verknüpft Polygone darstellen. Dreidimensionale Objekte werden durch Netze verschiedener Polygone (Drei- oder Vierecke) angenähert. Hierbei wird die Oberfläche des Ergebnisses umso glatter, je kleiner die zusammengesetzten Polygone sind. Man unterscheidet zwischen planaren und gekrümmten Polygonen. Planare Polygone (siehe Abbildung 4) bestehen aus Punkten und geraden Linien, gekrümmte Polygone (siehe Abbildung 5) erhält man, wie der Name schon sagt, durch gekrümmte Linien. In dieser Arbeit werden nur planare Polygone in Betracht gezogen, da diese sich für schnelle Annäherungsberechnungen mit linearen Gleichungen gut eignen. Die Genauigkeit der Ergebnisse dabei lässt sich beliebig einstellen. Die detaillierte Annäherung von gekrümmten Objekten mit planaren Polygonen erfordert zwar höheren Speicherbedarf als mit gekrümmten Polygonen, aber da Gebäude zum großen Teil lineare Umrisse aufweisen, hält sich der Speicherbedarf hier in Grenzen. Eine weitere Möglichkeit, 3D-Objekte in der Computergrafik zu beschreiben, heißt Voxel-Darstellung (Abbildung 6). Was Pixel im zweidimensionalen Raum sind, sind Voxel im dreidimensionalen Raum. Mit Voxel lassen sich dichte Körper oder Nebeldarstellungen realisieren. Voxel haben einen höheren Speicherbedarf als Polygone, da jeder Punkt gespeichert wird. Da in diesem Projekt nur die Hüllen der Gebäude rekonstruiert werden, werden hier keine Voxel verwendet.

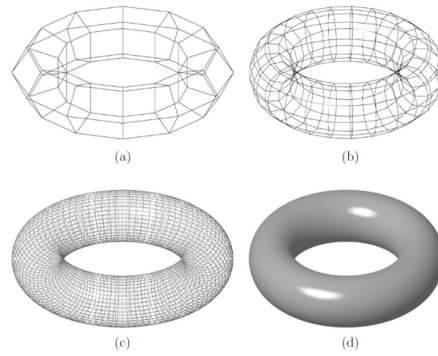


Abb. 4. Planare Polygone in verschiedenen Detailstufen. (Quelle: [3])

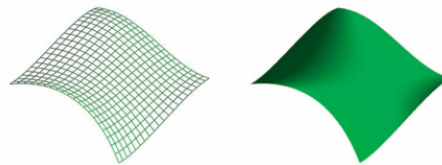


Abb. 5. Gekrümmte Polygone. (Quelle: [3])

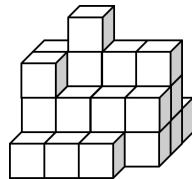


Abb. 6. Beispielhafte Darstellung von Voxel.

4 Möglichkeiten zur Umsetzung

Für die beispiel-basierte Rekonstruktion von Objekten gibt es bereits verschiedene Ansätze. Hier werden zwei Möglichkeiten gezeigt, die sich in einer ähnlichen Form für dieses Projekt eignen würden.

4.1 Analysis by Synthesis: 3D Object Recognition by Object Reconstruction

Grundlage für diesen Abschnitt ist das gleichnamige Paper der Wissenschaftler Mohsen Hejrati und Deva Ramanan aus der University of California, Irvine[2]. In diesem Ansatz gibt es bereits eine Datenbank mit verschiedenen geometrischen Formen (insgesamt über 1000 Modelle für Quader und Autos). Durch eine brute-force Methode vergleicht der Algorithmus die Formen des Eingangsmodells mit der Datenbank und erstellt mögliche Rekonstruktionen des 3D-Objekts. Das beste bzw. ähnlichste Ergebnis wird zurück gegeben (Beispiele siehe Abbildung 7). Die Herausforderung bei diesem Vorgehen ist die Rekonstruktion der vollen Komplexität eines Objekts. Außerdem können verschiedene Rekonstruktionen mehrere ähnliche Ergebnisse liefern, weswegen man sich hier nicht vollständig auf den Algorithmus verlassen kann, sondern den Benutzer einbinden muss. Der Benutzer kann sich dabei zum Beispiel die Anzahl der ähnlichsten Ergebnisse liefern lassen.

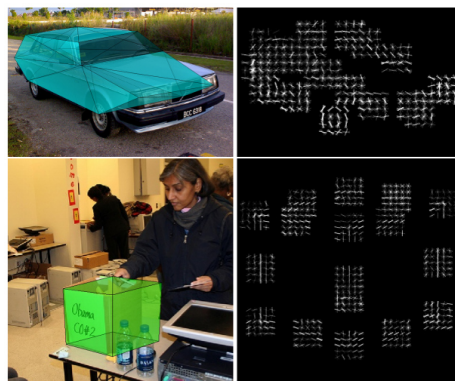


Abb. 7. Beispiele Analysis by Synthesis: Links: Eingabebild mit Rekonstruktion in blau bzw. grün darüber gelegt. Rechts: 2D-Templates, mit denen die Eingabebilder verglichen wurden. (Quelle: [2])

Vorgehen

Synthese Modell

Zunächst wird die Datenbank mit 3D-Modellen zum späteren Vergleich erstellt (Abbildung 8). Diese werden durch Punkte im dreidimensionalen euklidischen Raum dargestellt. Um perspektivisch veränderte Objekte wieder zu erkennen, gibt es weitere Parameter, um das Modell zu drehen, verzerren und zu projizieren. Für das hier behandelte Paper haben die Wissenschaftler sich für das *SUN primitive dataset*¹ und das *UCI-Car dataset*² entschieden.

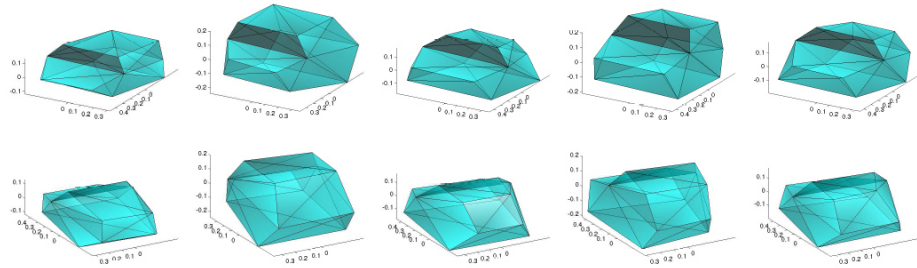


Abb. 8. Datenbank mit 3D-Modellen. (Quelle: [2])

Template Modell

Da ein 3D-Modell nicht direkt mit einem 2D-Bild verglichen werden kann, müssen hier zweidimensionale Template-Modelle aus den 3D-Modellen berechnet werden (Abbildung 9). (Quelle: [2])

Resultat

Für das Resultat werden die Template-Modelle mit Hilfe eines Inferenz-Algorithmus mit dem Eingabe-Bild verglichen und bewertet. Das oder die ähnlichsten Modelle werden als dreidimensionales Ergebnis geliefert.

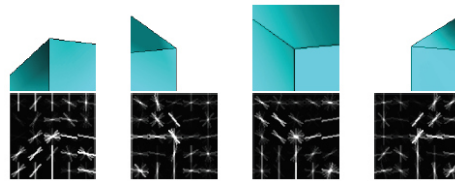


Abb. 9. Oben: Ecken aus Synthese Modellen, unten: Template Modelle. (Quelle: [2])

¹ SUN primitive dataset der Princeton Universität: <http://vision.princeton.edu/projects/2012/SUNprimitive/>

² UCI-Car dataset aus dem UCI Machine Learning Repository: <https://archive.ics.uci.edu/ml/datasets/Automobile>

Resultat

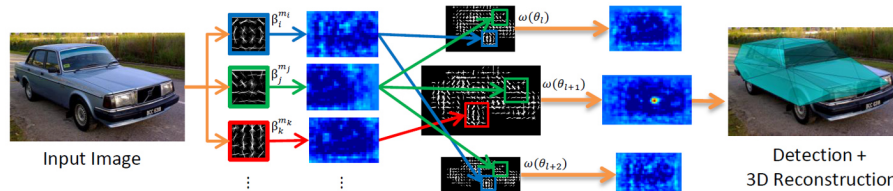


Abb. 10. Das Eingabebild wird mit den vorhandenen Template Modellen verglichen. Wenn mehrere dem Ergebnis ähneln, wird das ähnlichste Resultat als Ergebnis geliefert. (Quelle: [2])

4.2 Example-based 3D Object Reconstruction from Line Drawings

Bei diesem Verfahren, das von drei Wissenschaftlern aus China und Hong Kong[5] entwickelt wurde, werden die 3D-Objekte aus gegebenen zweidimensionalen Linien mithilfe einer gegebenen Datenbank aus 3D-Linienobjekten rekonstruiert. Dabei werden komplexere Objekte in mehrere Einzelteile zerlegt, die zunächst unabhängig voneinander rekonstruiert und später zum Resultat zusammengesetzt werden (Beispiel siehe Abbildung 11). Herausforderungen hierbei ist die Frage, welche der Linien in dem gegebenen Konstrukt perspektivisch nach vorne und welche nach hinten gehören. Die möglichen Lösungen beziehen den Benutzer mit ein. Diese wären:

- Per Gestensteuerung kann der Benutzer die Zusammenhänge zwischen den Linien zeigen
- Der Benutzer zeigt (durch Klick oder Geste) parallele Linien und rechtwinklige Ecken an

Dabei müssen Nachteile bei den interaktiven Lösungen berücksichtigt werden:

- Der Benutzer muss technikaffin sein, um mit dem System umzugehen,
- Es entsteht ein größerer Aufwand bei komplexen Objekten.

Die Planarität der Kanten und die Sichtbarkeit aller Ecken sind Voraussetzungen für das Verfahren in dem Paper. Außerdem müssen größere 3D-Objekte, die rekonstruiert werden sollen, in kleinere, weniger komplexe Einzelteile auseinander genommen werden können. In diesem Verfahren werden die genannten Einzelteile mit Objekten aus der Datenbank verglichen und die passenden 3D-Objekte für jedes identifiziert. Die Auswahl aus der Datenbank erfolgt mittels Maximum-a-posteriori Annäherung. Zuletzt werden die gefundenen 3D-Einzelteile aneinander gefügt.

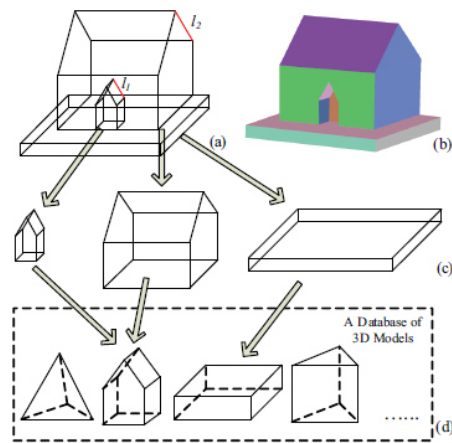


Abb. 11. Beispiel: das Eingabebild (a) wird in mehrere Einzelteile zerlegt (c), die unabhängig voneinander mit Objekten aus der Datenbank (d) verglichen werden. Die passenden Objekte werden zum Schluss zusammengefügt und das Resultat (b) zurück gegeben. (Quelle: [5])

3D Modelle in der Datenbank

Die 3D Modelle in der Datenbank, mit denen die Eingabeobjekte verglichen werden, werden manuell erstellt. Die Eckpunkte der Modelle sind Vektoren in einem euklidischen dreidimensionalen Koordinatensystem (Abbildung 12). Ein ungerichteter Graph zeigt an, welche Vertices miteinander verbunden sind.

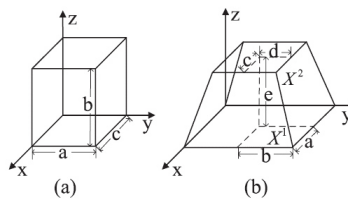


Abb. 12. Beispielhafte Darstellung eines Würfels (a) und eines Quaders (b) in einem 3D-Koordinatensystem. (Quelle: [5])

Vorgehen

1 Line drawing decomposition: Das Eingabeobjekt besteht aus einem Gerüst aus Linien. Die Linien zeigen die Kanten des Objekts im Bild an. Das Linienobjekt wird im ersten Schritt in mehrere, weniger komplexe Objekte aufgeteilt.

2 *Candidate 3D model generation:* Für jedes Einzelteil werden aus der Datenbank Kandidaten herausgesucht, die am ehesten in Frage kommen.

3 *3D reconstruction:* Aus den Kandidaten werden für die Einzelteile diejenigen selektiert, die am besten passen bzw. am ähnlichsten sind.

4 *3D part combination:* Die Kandidaten der Einzelteile werden zum großen Modell zusammengesetzt.

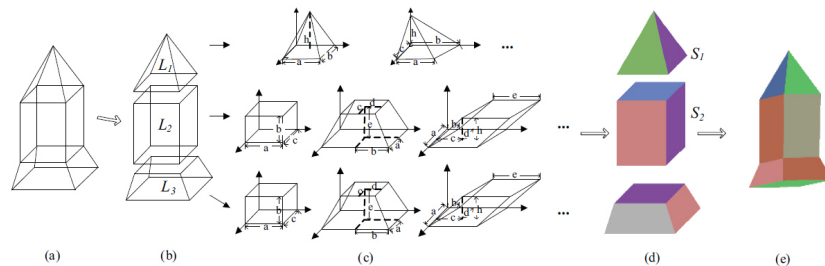


Abb. 13. Der Prozess der beispielbasierten Rekonstruktion: das Eingabemodell (a) wird zu weniger komplexen Einzelteilen (b) auseinander genommen. Diese werden mit der Datenbank (c) abgeglichen. Die Resultate (d) werden zum Ergebnis (e) zusammengefügt. (Quelle: [5])

5 Aktuelle Forschung, Forschungsgruppen und Konferenzen

Die Forschungsgruppe des Multimedia Lab aus dem Department of Information Engineering der Chinese University of Hong Kong hat ein Projekt mit dem Namen „Line Drawing“³, aus dem das Paper *Example-based 3D Object Reconstruction from Line Drawings* stammt. Sowohl vergangene als auch zukünftige Veröffentlichungen dieser Gruppe können für dieses Masterthema interessant sein. Eine jährlich stattfindende Konferenz, die für dieses Thema auch im Auge behalten werden soll, ist die IEEE Conference on Computer Vision and Pattern Recognition⁴. Dort werden nicht nur Vorträge zur Pattern Recognition, sondern in der Vergangenheit auch zum Thema Rekonstruktion von Objekten gehalten. Konferenzen, die auch ab und zu Vorträge zum Thema Objekterkennung und -rekonstruktion enthalten, sind SIGGRAPH (Special Interest Group on Graphics

³ Projekt „Line Drawing“ des Multimedia Lab, Department of Information Engineering, The Chinese University of Hong Kong http://mmlab.ie.cuhk.edu.hk/project_3d_line_drawing.html

⁴ IEEE Conference on Computer Vision and Pattern Recognition 2016: <http://cvpr2016.thecvf.com/>

and Interactive Techniques)⁵ und BMVC (British Machine Vision Conference)⁶. Diese finden auch jährlich statt.

6 Arbeitsplanung

Ähnlich wie in den beiden vorgestellten Papers soll in dieser Arbeit auch eine Datenbank aus parametrisierbaren Bausteinen wie etwa Quader oder Pyramiden verwendet werden. Die Parameter der Bausteine bestehen aus:

- Eckpunkte des Bausteins in x, y und z-Richtung
- Verschiebung des Bausteins t_x, t_y, t_z
- Parameter zur Rotation: Euler-Winkel oder Quaternionen

Anders als im „Analysis by Synthesis“ Paper sollen in dieser Arbeit nicht die Strukturen an den Objekten, sondern explizit die Umrisse der Objekte per Bilderkennung identifiziert werden. Somit erhält man eine Repräsentation des Objekts aus Linien. Mit diesen Linien können Vertices und Winkel erkannt werden, mithilfe derer man das Objekt mit den Bausteinen aus der Datenbank vergleichen kann.

Zunächst muss das Objekt aus dem Eingabebild extrahiert werden. Dann müssen daraus die Umrisse des Gebäudes extrahiert werden. Mit dem Ansatz des Analysis-by-Synthesis wird das 3D-Objekt dann rekonstruiert.

Projekt 1

In Projekt 1 soll die Umsetzung im zweidimensionalen Bereich erfolgen. Das Eingabebild wird nur aus standard-2D Formen wie Recht- und Vierecken bestehen. Es wird eine Datenbank mit 2D-Formen für den Vergleich geben. Wie oben beschrieben werden hier für die Bausteine die Vertices für x und y eines Koordinatensystems gespeichert. Außerdem werden ungerichtete Graphen gespeichert, um zu zeigen, welche der Vertices miteinander verbunden sind. Außerdem soll es für den Satz von Parametern dieses Projekts eine Fehlerfunktion geben, die eine Ähnlichkeitsanalyse durchführt, damit das Ergebnis optimiert werden kann. Diese Optimierung bzw. die Minimierung der Fehlerfunktion soll schon in diesem Projekt erfolgen. Dieselbe Optimierung soll später im 3D Modell auch weiter verwendet werden können. Außerdem soll in diesem Projekt schon mehrere Formen für die Verarbeitung miteinander kombinierbar sein.

Projekt 2

Projekt 2 wird eine Erweiterung des Projekts 1 von 2D auf 3D sein. Eine Projektionsmatrix soll für die perspektivische Abbildung der 3D Objekte hinzugefügt werden. Alle Features aus Projekt 1 sollen hier in 3D umgesetzt werden.

⁵ Special Interest Group on Graphics and Interactive Techniques: <http://www.siggraph.org/>

⁶ British Machine Vision Conference: <http://www.bmva.org/>

Masterarbeit

Die Masterarbeit wird die Erweiterung von Projekt 2 zu einem vollständigen System sein. Dieses System könnte eine Applikation für ein Smartphone/Tablet oder ein Programm auf dem PC sein. Die Bausteinfindung soll hier automatisiert und optimiert werden. Optional sollen Texturen aus dem Eingabebild extrahiert und in die 2D-Rekonstruktion hinzugefügt werden (für eine realistische Darstellung).

Umsetzung Vorgesehen ist eine Implementierung in Java. Damit wäre auch die Möglichkeit gegeben, aus dem Programm eine portable App für Android Smartphones oder Tablets zu gestalten. Frameworks und Bibliotheken, die für dieses Projekt relevant sein könnten, sind folgende:

- Das Computergrafik Framework der Computergrafik Forschungsgruppe von Prof. Philipp Jenke. Das Framework basiert auf Java und OpenGL und kann zur grafischen Darstellung der verschiedenen Objekte verwendet werden.
- OpenCV ist eine Computergrafik Bibliothek von Intel. Mithilfe von OpenCV lassen sich unter anderem Objekt- und Gestenerkennung simpel realisieren.
- Außerdem können verschiedene Geometrie Bibliotheken für benötigte mathematische Berechnungen eine Hilfe sein.

7 Fazit/Ausblick

Die korrekte und detailgenaue Objekterkennung ist ein essentieller Bestandteil dieses Projekts, sie bildet die Grundlage aller weiteren Schritte. Die Berechnungen zur Annäherung der Gebäudeumrisse müssen daher optimal sein.

Interessant könnte dieses Projekt für Immobilienmakler sein, die einem Kunden das Immobilienobjekt anhand eines 3D-Renderings zeigen möchten, ohne dass der Kunde vor Ort sein muss. Auch wäre der Einsatz bzw. die Einbindung in die Spieleindustrie möglich, zum Beispiel in Simulationsspielen, in denen man seine eigene Wohngegend nachbauen kann. Die Methodik der Objekterkennung und -rekonstruktion lässt sich selbstverständlich auch für andere Objekte als Gebäuden einsetzen, zum Beispiel zu Marketingzwecken, indem der Kunde mit seinem Smartphone Werbeplakate einscann, um das Produkt virtuell als 3D Modell zu sehen.

Literatur

1. Einfamilienhaus Rendering, <https://pixabay.com/de/einfamilienhaus-villa-rendering-1026372/>
2. Hejrati M., Ramanan, D.: Analysis by Synthesis: 3D Object Recognition by Object Reconstruction. IEEE Conference on Computer Vision and Pattern Recognition (2014)
3. Nischwitz, A., Fischer, M., Haberäcker, P., Socher, G.: Computergrafik und Bildverarbeitung, Band I: Computergrafik. 3. Auflage (2011)

4. Robin, P.: How Apple Creates 3D Flyover Maps, <http://www.applegazette.com/apple-inc/how-apple-creates-3d-flyover-maps/> (2012)
5. Xue, T., Liu, J., Tang, X., Example-Based 3D Object Reconstruction from Line Drawings. IEEE Conference on Computer Vision and Pattern Recognition (2012)