

Objektrepräsentation und Merkmalsextraktion

Objektrepräsentation

- Die Repräsentation und Anpassungsphase sind entscheidend für den Nutzen eines Wiedererkennungssystem
- Objektrepräsentation vereinfacht komplizierte und aufwendige Voll-3D-Darstellungen
→ Kompakte Form und effiziente Speicherung
- Zu beachten ist, dass genügend Information erhalten wird
→ verrauschte oder verdeckte Darstellung müssen immer noch zum gespeicherten Modellen passen
- In der Anpassungsphase werden Techniken verwendet die eine umfassende Suche, des in der Datenbank gespeichertem Modell vermeiden

Objektrepräsentation

- Alle gängigen Erkennungssysteme werden in zwei Gruppen der Objektrepräsentation klassifiziert
 - (1) Objekt-zentrierten (object-centred)
 - (2) Beobachter-zentrierten (viewer-centred)

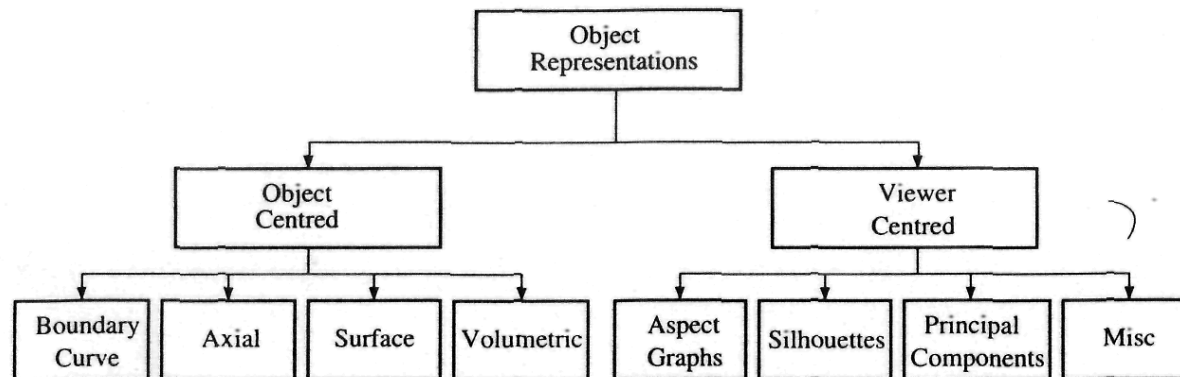


Fig. 4.2. A taxonomy of representations as applied to object recognition

Objektrepräsentation

- Unterteilung der objekt-zentrierten Repräsentation
 - (1) Begrenzungsflächenmodell und kurvenbasierende Repräsentation
 - (2) axiale Repräsentation
 - (3) Oberflächenbeschreibung
 - (4) Volumenmodell

1. Begrenzungsflächenmodell

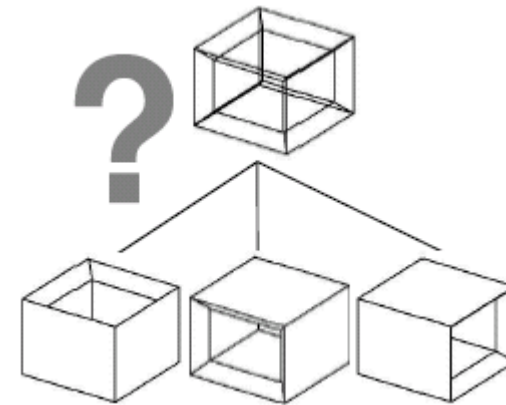
- Speicherung des Modells als Tripel:
 - (1) Satz von Oberflächen
 - (2) Satz von Raumkurven, die die Überschneidungen der im ersten Schritt gespeicherten Oberflächen repräsentieren
 - (3) Ein Graph, der die Oberflächenverbindungen (surface connectivity) und andere Informationen über die Oberfläche beinhaltet

Begrenzungsflächenmodell

- einige der frühesten Systeme:

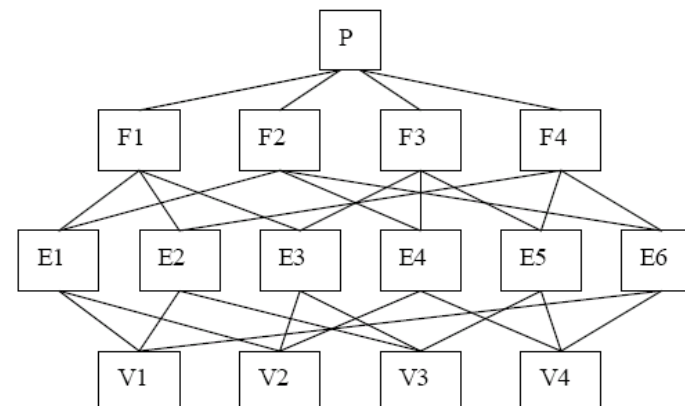
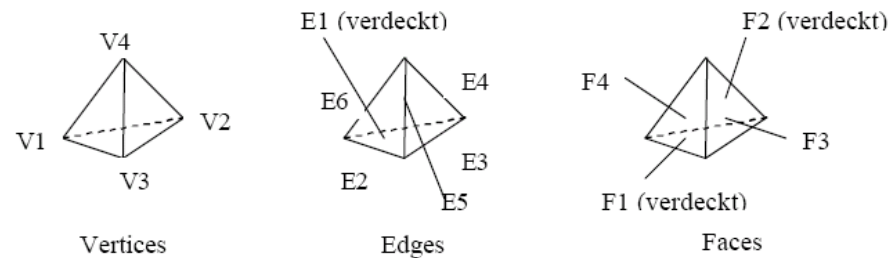
Drahtgittermodell

- 3D-Modell aus Verzeichnis möglicher Kantenverbindungen
- für Objekte aus ebenen Flächen und geraden Kanten



Flächen-Kanten-Punkt Repräsentation (Surface-Edge-Vertex)

- enthält Liste der Kanten und Eckpunkte eines Objektes
- Beschreibung der existierenden Oberflächenbeziehungen



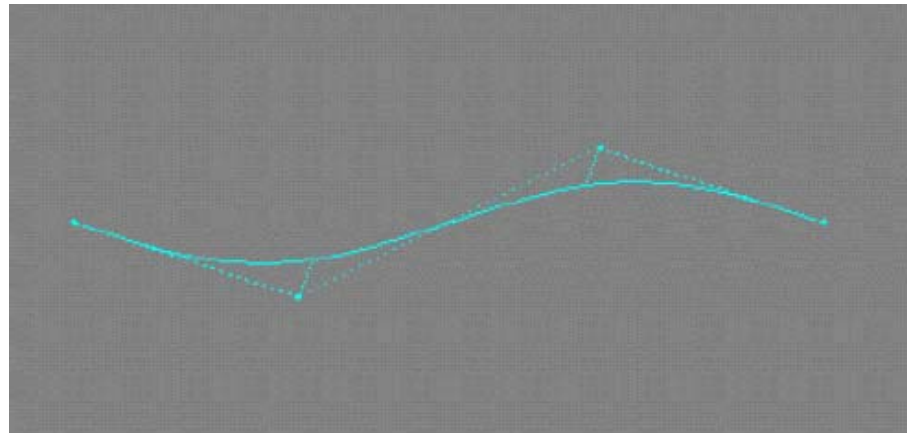
Anwendung

- Bhanu und Faugeras
 - verwendeten hierarchisch stochastische Kennzeichnungen zur Lösung des Segment matching Problems
 - Teil eines Objekts wird mit einer übergeordneten Modellrepräsentation gematched
 - Merkmale wie Länge der Segmente oder Winkel zwischen ihnen werden aus der Repräsentation extrahiert
 - Berechnung eines Vektors für jedes Modellsegment, bestehend aus der Wahrscheinlichkeit, dass dieses Segment des Modells einem jeweiligen Segment des Objektes entspricht
- Verwendung von Diskontinuitäts-Informationen
 - verwendet von Fan
 - Begrenzungsflächenmodell
 - Glättung von range images unter Verwendung von Gauss'schen Filtern
 - Berechnung von positiven und negativen Kurvenextremwerten und des Nullpunkt der größten Krümmung
 - Verbindung der erhaltenen Werte um Vorsprünge und Vertiefungen zu identifizieren
 - Durchführung von spatial linking um abgegrenzte Flächen zu vermindern



- quadratische Oberflächenfunktionen werden an segmentierte Oberflächen angelegt
 - Information wird in einem zugeordneten Graph repräsentiert
 - abgegrenzte Flächen werden ihren Charakteristika nach gekennzeichnet
 - Informationen werden in einer Reihe von Knoten gespeichert wie sichtbarer Bereich, Orientierung und Schwerpunkt für jedes Oberflächenstück
- Verwendung von Invarianz
 - kurvenbasierende Repräsentation
 - Invariante sind Bestandteile von geometrischen Konfigurationen, die unter Transformation unverändert bleiben
 - Schnittpunkte, Kollinearität und Tangenten bleiben unter Projektion erhalten
 - kann auch bei kleineren Deformierungen von Objekten angewandt werden
- System: LEWIS (Zisserman u.a.)
- Kanten werden aus einem eingegebenen Bild extrahiert
 - Berechnung von drei verschiedene Arten von projektiven Invarianten
 - invarianten Werte werden in Hashtabelle eingetragen
 - Satz der möglichen Interpretationen wird einem Verifikationsprozess unterworfen um das passende Modell zu finden
 - System MORSE, Erweiterung auf 3D Strukturen durch Auferlegung von Einschränkungen für die Oberflächen

- Spline im besonderen B-Spline
 - kurvenbasierte Repräsentationen beziehen oft Kurven mit besonderen Eigenschaften ein
 - Kurven geben unverwechselbare Eigenschaften der Oberfläche wieder
 - Spline: Kurve, die durch eine zusammenhängende Kette von polynomialen Funktionen und Stützpunkte definiert wird. B-Splines sind spezielle Splines



B-spline

→ mathematische Darstellung

$$s(u) = [x(u), y(u), z(u)] = \sum_{i=-k}^g c_i B_{i,k+1}(u)$$

u – Parameter, definiert entlang der Bogenlänge der Kurve

c_i - sind die Kontrollpunkte

k – bezeichnet Verwendung von k-gradigen Polynomen

g – Anzahl der inneren Knoten

$B_{i,k+1}$ - B-Spline Basisfunktion

Eigenschaften der Spline mit besonderer Bedeutung für die Objekterkennung

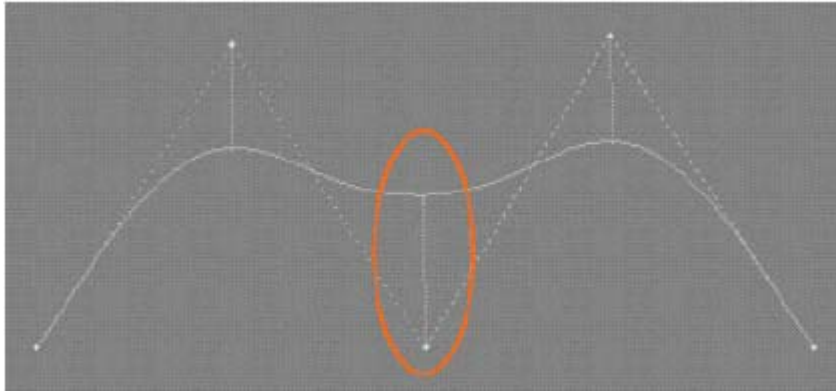
- Eine einfache Repräsentation in der abgeleitete Information analytisch aus der Repräsentation abgeleitet werden können. Diese Eigenschaft ermöglicht die schnelle Evaluation bei Krümmung oder Drehung an Punkten entlang der Oberfläche
- Die Repräsentation ist unveränderlich sowohl bei affiner als auch bei projektiver Transformation. Dies ist wichtig im Erkennungsprozess, da nicht viel Nutzen in einer Repräsentation besteht, die sich mit einer Drehung oder Projektion des Objektes verändert.
- Splines haben eine fortpflanzungsfähige Natur, was bedeutet, dass eine vorgegebene Spline-Kurve in einem anderen Muster oder Skale wiederverwendet werden kann. Dies ist nützlich beim Anpassen von Kurven, wenn sie weiter verkleinert werden um eine genauere Einschätzung der Position zu ermöglichen.

Eigenschaften der Spline mit besonderer Bedeutung für die Objekterkennung (2)

- Splines sind gleichzeitig eine lokale und eine globale Repräsentation. Sie sind global, da die Summe der einzelnen Polynome die gesamte Kurve definieren. Trotzdem bieten sie auch eine lokale Kontrollierbarkeit an, da die Bewegung eines Kontrollpunktes die Form der Kurve nur in der unmittelbaren Umgebung des Kontrollpunktes verändert.
- Sich annähernde B-splines sind robuster als interpolierte Splines in verrauschter Umgebungen, da sich die Kurve über schnelle Fluktuationen glättet.

Anwendung

- Gueziec und Ayache
 - anpassen der Gipfelinien aus medizinischen Bildern an NURBS Kurven (Nonuniform Rational B-Spline)



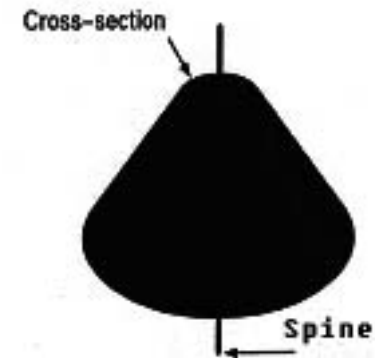
- jeder Kontrollpunkt kann gewichtet werden, ungleichmäßige Knotenabstände
- schätzen der Spline-Knoten durch Optimierungsfunktion basierend auf den Normalenvektoren der Oberfläche

Hauptprobleme

- Begrenzungsflächenmodell
 - oft unzureichend
 - Informationsverlust durch Reduzierung von Oberflächen auf Kurven
 - Vernachlässigung von Anwendungsgebieten
- kurvenbasierende Repräsentation
 - viele Techniken nicht robust genug gegenüber Störungen (Verauschen)
 - beeinträchtigt zuverlässige Merkmalsextraktion und Genauigkeit
- Gaussch'e Glättung
 - verlangt Einschätzung von Parametern wie Fenstergröße und Rauschvarianz
- Spline
 - trotz recht komplizierter Kurven Schwierigkeiten beim matching der Repräsentation aufgrund des Knotenproblems
 - Knotenproblem: mit unterschiedlicher Platzierung der Knoten lässt sich ein neuer entsprechender Satz Kontrollpunkte ableiten
 - weitere Stufe der Merkmalsextraktion notwendig
 - Einbußen der Effizienz auf Kosten der Genauigkeit

2. Axiale Repräsentation

- bekanntes Beispiel der generalisierte Kegel
 - besteht aus zwei Komponenten
 - eine Raumkurve (spine)
 - Querschnittsform die sich entlang der Achse sanft verformt
 - Zerlegung in zwei Darstellungen
 - Ein Strich als Projektion entlang der Raumkurve
 - Ellipse als Teil des Querschnitts
- Anwendung im System ACRONYM



Anwendung

- Das u.a.
 - Algorithmus der hierarchische CAD-Modelle für verschiedene Objekterkennung verwendet
 - System hängt von Mehrfachlösungen oder verschiedenen CAD-Modellen ab
 - Darstellungsbeschreibung für jedes Teilstück unter Verwendung von Vielecken, konvex/konkav Kanten, Kurvenextremas, Oberflächennormalen und generalisierten Zylindern
 - Merkmalsextraktion aus den Objektteilen
 - Matching beginnt mit der Identifizierung einer Region am allgemeinsten CAD-Modell
 - immer spezifischer werdender Ablauf der Merkmalsextraktionsschritte für eine feinere Erkennung des Objektes

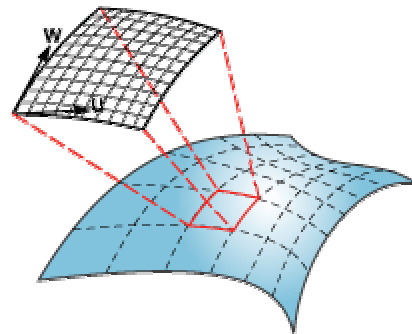
Probleme

- generalisierter Kegel exzellent geeignet um gestreckte Objekte mit axialer Geometrie darzustellen
 - in Natur kommen aber meist komplexere Objekte vor
 - Deformationsparameter entlang des Spine können durch Fehler von Verrauschungen beeinflusst werden
- Daher allseits eingeschränktes Anwendungsfeld

3. Oberflächenbeschreibung

- ursprünglich galt die Aufmerksamkeit der Forscher der Verwendung von Vielecken und im besonderen quadratische Oberflächen
- diese wurden den segmentierten Teilen der Oberflächen angeglichen
- keine Extraktion von hervorstehenden Oberflächen
- Oberflächenunterteilung in patches

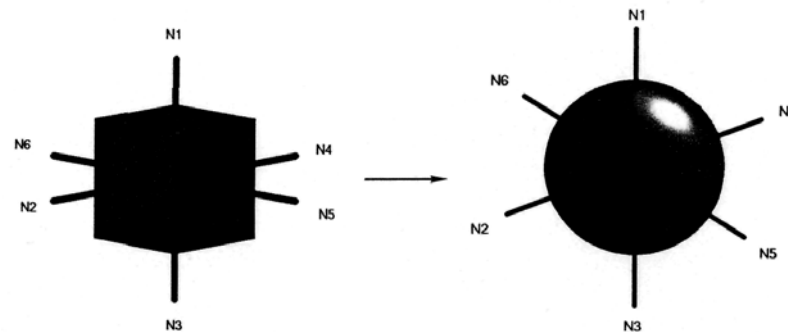
→ patch



- Mängel bedürfen verschiedener Merkmalsextraktion- und Segmentierungstechniken

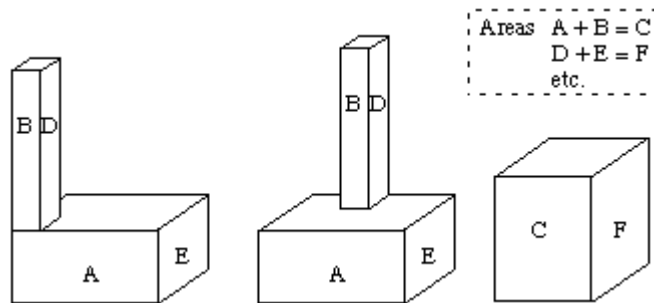
Extended Gaussian Image und Complex Extended Gaussian Image

- EGI und CEGI
- Kernaussage diese Techniken ist die Darstellung eines 3D Objektes aus der Sammlung der Oberflächennormalen
- der Satz der Oberflächennormalen kann in dem Gaussch'en Bereich dargestellt werden
 - Fixierung der Vektorenden in der Mitte des Bereichs
 - Spitzen der Normalen durchstoßen den Verbindungsbereich in passender Ausrichtung



- ursprüngliche GI hatten den Nachteil das es nicht umkehrbar in plane Objekte war
- Entwicklung des EGI mit gewichteten Normalenvektoren der Oberflächen und die Einbeziehung von gewölbten Objekten durch Abbildung der Umkehrung der Oberflächenwölbung

Problem beim EGI:



→ verschiedene Objekte haben gleiches EGI

- Entwicklung des CEGI dem Normalendistanzen hinzugefügt wurden

Probleme

- Darstellungen sind nicht informationserhaltend (Ausnahme CEGI)
- nicht einzigartig
- Schwierigkeiten bei Szenen mit mehreren Objekten die sich gegenseitig verdecken

Segmentierung der Oberfläche in gesonderte Teilstücke (feature patch)

- Besl und Jain
 - Verwendung der Gauss'schen Krümmung und der mittleren Krümmung (mean curvature)
 - Oberflächen Segmentierung in acht fundamentale patch-Typen

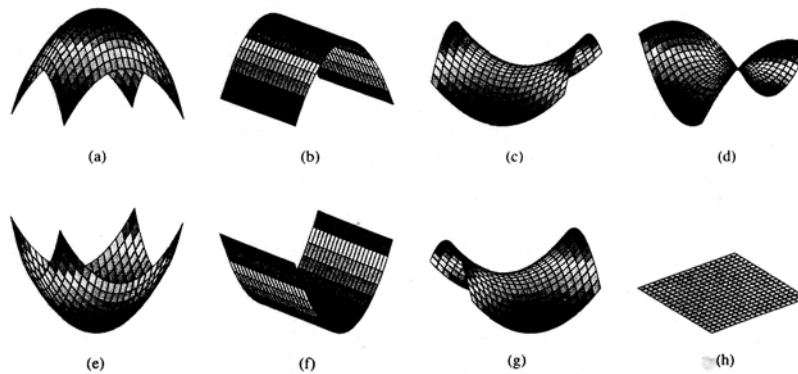


Fig. 4.8. (a) Peak: $H < 0$ and $K > 0$; (b) Ridge: $H < 0$ and $K = 0$; (c) Saddle Ridge: $H < 0$ and $K < 0$; (d) Minimal: $H = 0$ and $K < 0$; (e) Pit: $H > 0$ and $K > 0$; (f) Valley: $H > 0$ and $K = 0$; (g) Saddle Valley: $H > 0$ and $K < 0$; (h) Planar: $H = 0$ and $K = 0$

- zusätzliche Extrahierung von Sprungstellen und Extrempunkten
- Verbesserung der Segmentierung durch einen Anpassungs- und Wachstumsalgorithmus
- Verbindung der angenäherten Oberflächenregionen zu einer gemeinsamen Oberflächenbeschreibung

Probleme

- Oberflächenpatch-bezogene Repräsentationen haben ähnliche Schwierigkeiten wie die Systeme die auf dem Begrenzungsflächenmodell basieren
- viele Systeme nicht robust genug
- anfällig gegenüber Verdeckungen und Verauschungen
- viele Darstellungen sind nicht einzigartig
- Techniken sind rechnerisch aufwendig und verlangen komplexe Kontrolltechniken

Weitere Anwendungen

- Spline-bezogene Repräsentationen
 - angewandt in der Erkennung von Freiform-Objekten
 - Freiform Objekte bestehen aus Kurven oder Flächen, die mit Hilfe von Stützpunkten und Randbedingungen (z.B. Stetigkeiten) definiert werden. Sie werden mathematisch durch Polynome dargestellt
 - Beispiel: Spline, NURBS
 - 2D-Snake-Theorie verwendet von Liao und Medioni
 - internale und externale Kräfte wirken auf ursprünglich einfache Oberflächen ein um sie an die 3D-Datenpunkte anzupassen
 - Anpassung quadratischer Spline durch Anwendung iterativer Energie minimierender Algorithmen

→ Darstellung durch Freiform Objekte nicht weit verbreitet aufgrund der Schwierigkeiten bei der automatischen Oberflächenanpassung an Entfernungsdaten und Problemen von parametrischen Repräsentationen von Objekten

Extraktion von Merkmalen abgeleitet aus der differentialen Geometrie

- (1) globale Repräsentation
 - (2) lokale Repräsentation
 - (3) hybride Repräsentation
- die globale Repräsentation verwendet Merkmale und Parameter um das gesamte Volumen eines Objektes zu beschreiben
→ Beispiel: Simplex angle image und das aufblasbare Ballon-Modell (inflation baloon)

Simplex angle image (SAI)

- Netz das nahe an den Datenpunkten anliegt wird erstellt
- äußere und innere Kräfte werden an die Knoten angelegt und schrittweise verfeinert bis ein weiches Netz entsteht
- weitere Kräfte werden angelegt um eine lokale Gleichmäßigkeit zu erreichen
- nach Fertigstellung des Netzes wird eine Gauss'sche Sphäre angelegt mit identischer Anzahl an Knoten
- die der simplex angle entsprechenden Knoten werden in der Sphäre gespeichert
- dieses kugelförmige Netz wird als Simplex angle image bezeichnet

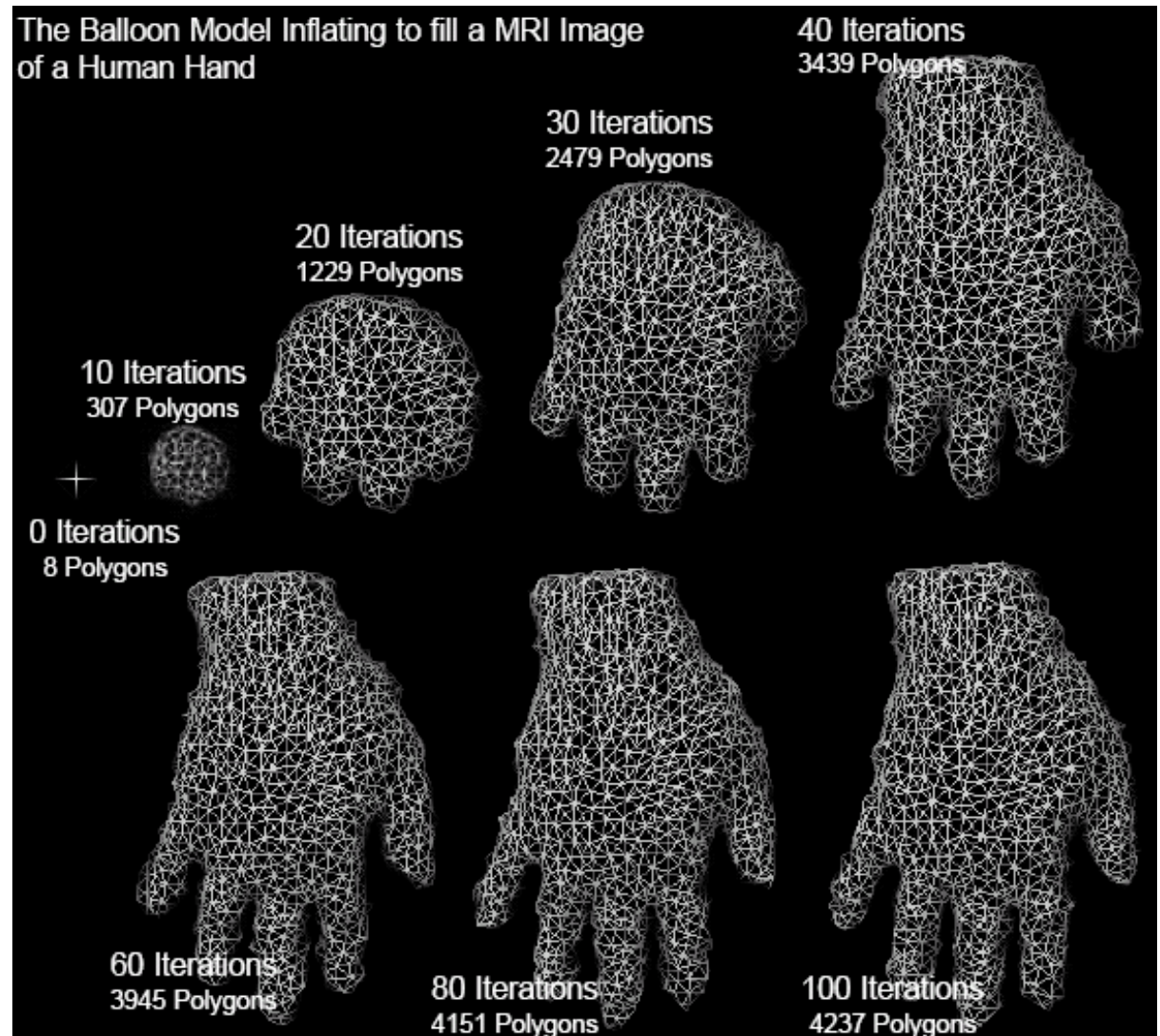
→ an sich robust aber nur verwendbar für Objekte die topologisch ähnliche einer Sphäre sind, außerdem Speicherung eines Netze für jedes Objekt damit ein effizientes matching ablaufen kann

Aufblasbarer Ballon-Modell

- Beginn mit einem engenmaschigem dreieckigem Netz im innern des Objektes
- Ballon wird „aufgeblasen“
- bis alle Dreiecke des Netzes mit der Objektoberfläche übereinstimmen

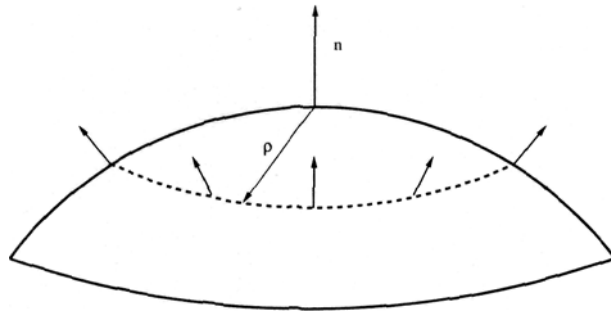
→ Nachteil

- keine Möglichkeit der Selbstunterbrechung
- Entwicklung einer Art Netzregelung erforderlich



TOSS System

- für lokale strukturelle Beschreibungen
- macht Gebrauch von zwei extrahierten Merkmalen eines Objektes
- ein Satz von 3D-Kurven zur Darstellung der Diskontinuität in den Vielecken
- Winkel zwischen den „Super Segmenten“ welche aus den Annäherung der Kurven gruppiert worden
- für Freiform Objekte wird eine splash-Repräsentation verwendet



- diese Repräsentation findet einen Bereich mit hoher Krümmung im Objekt und berechnet eine Oberflächennormale n
- eine kreisförmige Scheibe mit den Radius ρ wird definiert
- danach werden Oberflächennormalen abstandsweise auf dem Kreis berechnet
- ein Satz von „splashes“ bei variierendem ρ rund um die ursprüngliche Oberflächennormale berechnet heißt „super splashes“
- wenn Repräsentation berechnet, werden Hash-Tabellen basierende Strategien zum Erkennen herangezogen

Problem

- Repräsentation ist sensitiv gegenüber Störungen die die Kontur auf Oberflächen verändern
- dadurch Beeinflussung von Merkmalen die aus Krümmung oder Oberflächennormalen abgeleitet werden
- Beeinträchtigung von Genauigkeit und Robustheit
- bei globalen Systemen Schwierigkeiten mit Überdeckungen
- lokale Systeme langatmig und stör anfällig

→ daher beinhalten neuere Erkennungssysteme globale und lokale Darstellungen
Bsp.: spin images und COSMOS

Spin Images

- von Johnson
- entwickelt um 3D-Oberflächen zu matchen
- dieses Schema nimmt an, dass die Oberfläche ursprünglich als vieleckige Oberflächennetze repräsentiert werden
- es werden Orientierungspunkte O_i (ein Punkte mit entsprechendem Normalenvektor) und spin images S_i kreiert
- spin image ist ein 2D Histogramm der Oberflächenstellen, die von einer Ebene geschnitten werden, die sich um die Normale eines Orientierungspunkt drehen
- jeder Orientierungspunkte auf der Oberfläche des Objektes hat eine einzigartige spin map
 - spin map daher das die polarentgegengesetzten Winkel nicht in einem zylindrischem Koordinatensystem berechnet werden können und die Basis kann sich um die Achse drehen, ohne einen Effekt auf die Koordinaten der Punkte im Bezug auf die Basis zu haben
- spin images sind unveränderlich gegenüber strengen Transformation, da sich die Position der Punkte auf einer Oberfläche mit Bezug auf den Orientierungspunkt beschreiben

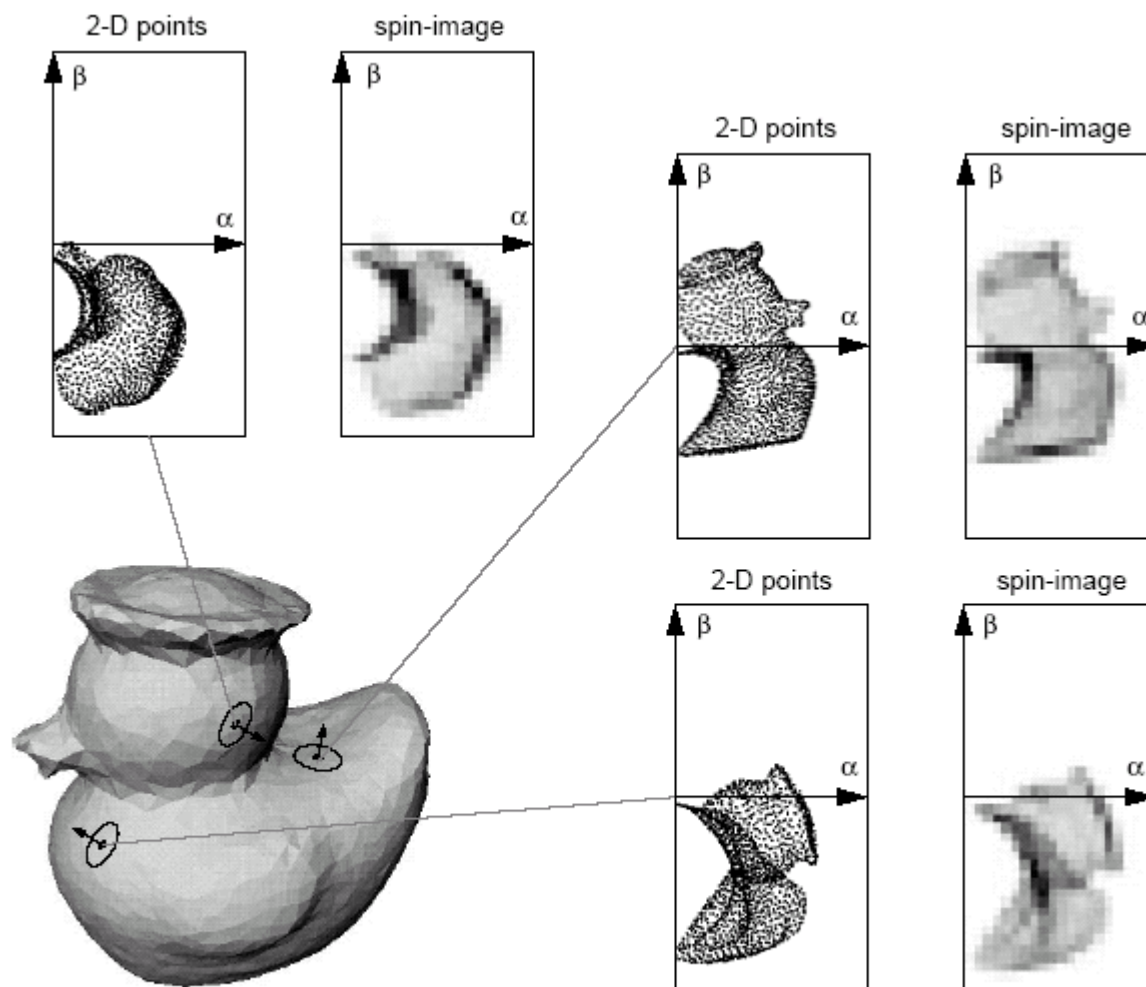


Figure 2-2: Spin-images for three oriented points on the surface of a rubber duckie model. The 3-D position of vertices in the mesh are mapped into 2-D using the spin-map for each oriented point basis. By accumulating 2-D points in discrete bins, spin-images are generated.

COSMOS

- Formerkennungs-System
- entwickelt von Dorai und Jain
- Schema extrahiert spezifische geometrische Aspekte eines Objektes
- wird definiert durch Verwendung von Differentialgeometrie bezogenen Konzepten
- System nimmt an das die Oberfläche S zumindest teilweise glatt ist und es zumindest zweifach differenzierbar ist
- infolgedessen können Quantitäten wie Orientierung und Krümmung berechnet werden
- als Mittel zur Formbeschreibung wird Shape spectral function(SSF) vorgestellt
 - SSF erlaubt die quantitative Charakterisierung der Formkategorien im Objekt
- die Orientierung von Teilstücken des Objektes werden in Terme ihrer durchschnittlichen Oberflächennormalen zusammengefasst
- globale Strukturen werden außerdem in den Informationen der Teilstücke codiert
- der Kernpunkt des COSMOS Systems ist die Definition von Constant shape maximal patches (CSMP)
 - Region wo Punkte den selben Formindex haben und achtfach miteinander verbunden sind

- wenn die Segmentierung in passende CSMP's durchgeführt wurde, werden
 - Orientierung der Oberfläche
 - Verbindungsinformationen am Objekt
 - hervorstechenden lokalen Oberflächenmerkmaleextrahiert
- es wird ein diskretes Form-Spektrum definiert welches als Histogramm dient um Bereiche mit der selben Form zusammenzufassen

→ Nachteil ist, dass mit dem Wachstum der Regionen in Verbindung mit Gebieten mit einem konstanten Formindex es zu Vermischungen zwischen den Regionen kommen kann

→ COSMOS System nicht getestet für vollständige 3D Objekte aus einer Vielzahl von Blickwinkeln

4. Volumenmodell

- Hauptdarstellungstechniken sind:

(1) Voxels

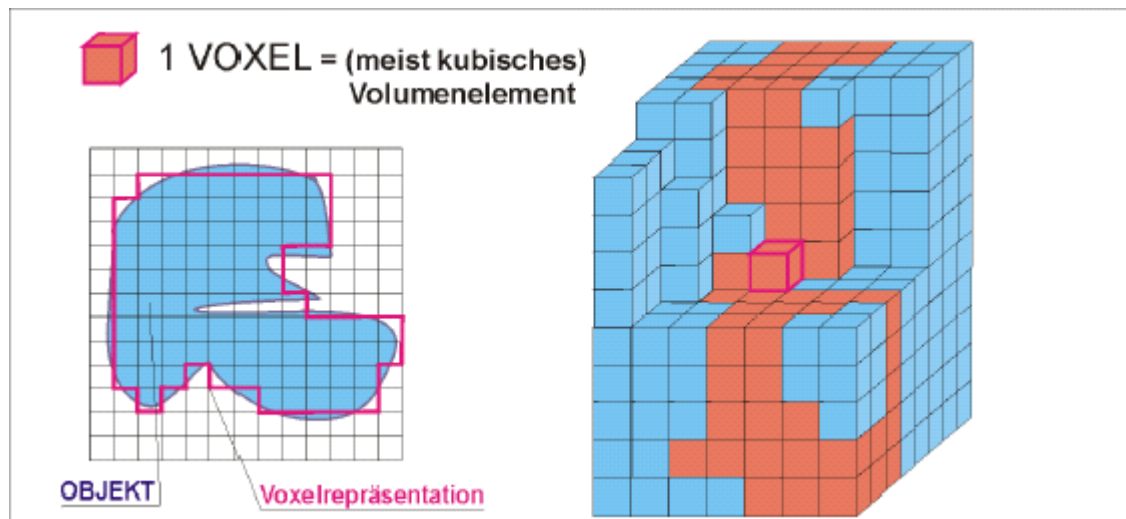
(2) 8-Bäume (octrees)

(3) Superquadriken

(4) Konstruktion mit Raumprimitiven (CSG – constructive Solid geometry)

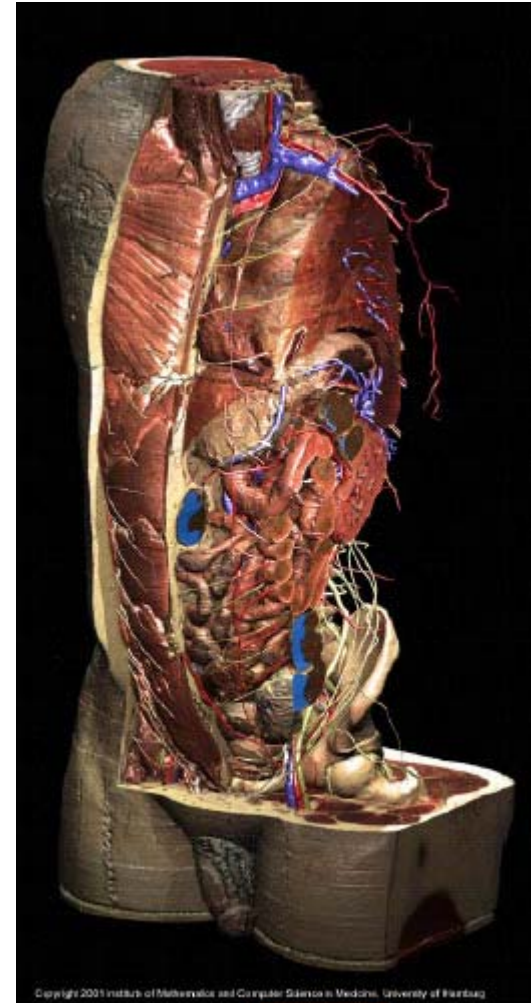
Voxel

- eine Methode zu Raumaufteilung
- ein Voxel ist die kubische Entsprechung eines Pixels
- Kantenlänge muss definiert werden
- je genauer die Objektbeschreibung sein soll desto kleiner muss eine Voxel gewählt werden
- Voxeldarstellungen sind sehr speicherintensiv da die Datenmengen für genauere Darstellungen recht groß werden kann
- einzelne Voxel können spezielle Eigenschaften annehmen wie Transparenz, Lichtintensität ect.



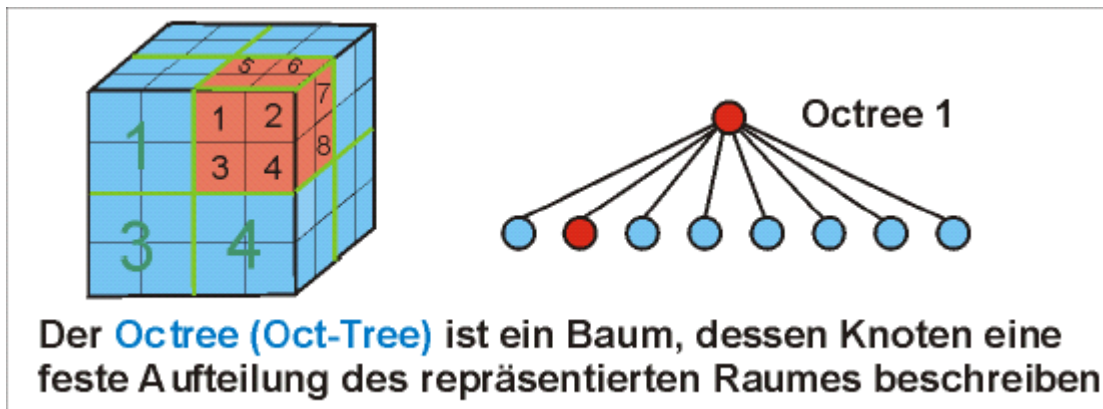
Voxel (2)

- Voxeldarstellungen finden zum Beispiel Anwendung in der Medizin
 - störenden Voxel können mit der Eigenschaft Transparenz belegt werden
 - Haut wird ausgeblendet
- Schnitte sind ebenfalls leicht realisierbar



8-Bäume

- Baumstruktur
- jeder Knoten entspricht einem gewissen Bereich im dreidimensionalen Raum
- Unterteilung in acht Volumenelemente mit 3 Möglichkeiten
 - voll im Objekt
 - leer
 - partiell im Objekt
- in den ersten beiden Fällen endet der Baum an dieser Stelle und Information kann gespeichert werden
- liegt das Volumenelement nur partiell im Objekt kann diese wieder in acht Teilstücke zerlegt werden
- nicht so Speicherintensiv wie die Voxeldarstellung aber im Verhältnis zu anderen Repräsentationen immer noch hoch



Superquadriken

- die Tatsache das ihre Beschreibungen immer geschlossene Oberflächen produzieren bedeutet, dass sie immer ein Volumen umschließen

$$S = a \begin{pmatrix} a_1 & C_u^{\varepsilon_1} & C_u^{\varepsilon_2} \\ a_1 & C_u^{\varepsilon_1} & S_u^{\varepsilon_2} \\ & a_3 & S_u^{\varepsilon_1} \end{pmatrix} \quad \text{where } -\frac{\pi}{2} \leq u \leq \frac{\pi}{2}, \quad -\pi \leq v \leq \pi$$

- Anwendung um Volumina anzupassen, globale Deformationen, Biegungen, Drehungen und boolesche Zwänge zu beschreiben

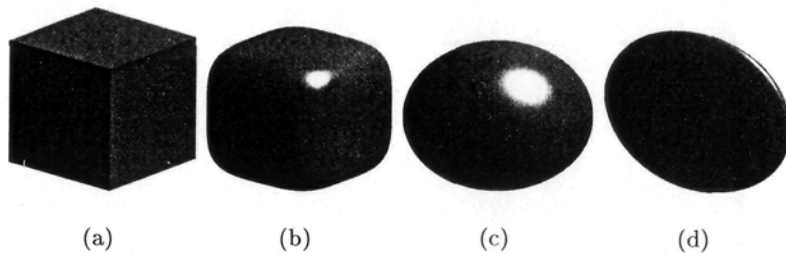
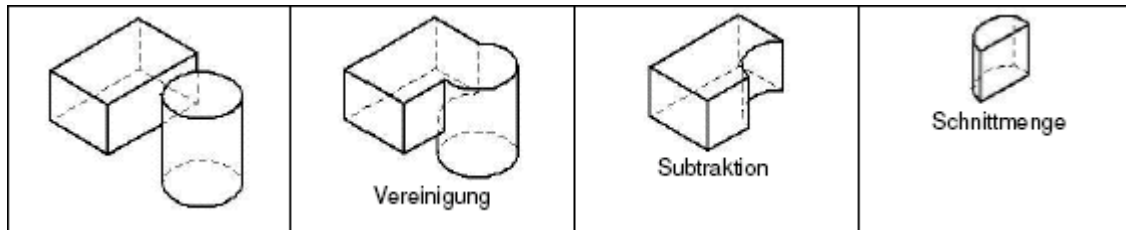
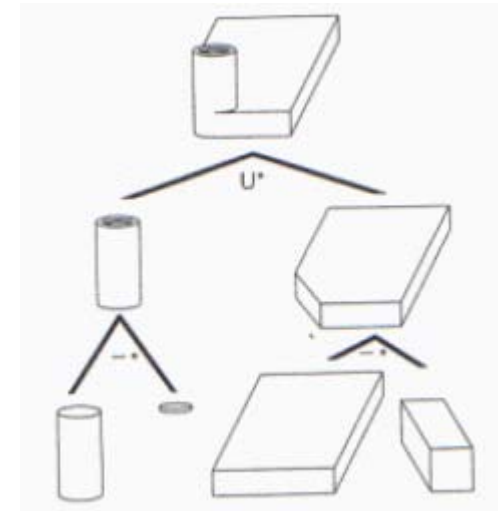


Fig. 4.11. The effects of the squareness and aspect ratio parameters on a superquadric's shape: (a) Squareness parameters set to 0; (b) Squareness parameters set to 0.5; (c) Squareness parameters set to 1; and (d) Aspect ratio set to (0.2,0,0)

Konstruktion mit Raumprimitiven

- Verknüpfung einfacher geometrischer Körper
- Körper wie Würfel, Kegel, Zylinder ...
- Verknüpfung erfolgt über boolesche Operationen wie Schnitt, Differenz, Vereinigung
- Entstehungsreihenfolge wird in einem binären Baum abgespeichert
- Vorteile:
 - leicht interaktiv bedienbar
 - Formen können leicht geändert werden
 - auch Entfernung von Material
 - kompakte Speichermöglichkeit
- Nachteile:
 - nicht immer eindeutige Darstellung
 - hohe Rechenzeit wegen vielen Punkten und Kanten notwendig



Quellenangabe:

1. „Representation and Recogniton of 3D Free-Form Objects“, G.Mamic und M.Bennamoun, 2001
2. „Rechnergenerierte Inszenierung von Animationen“, Axel Hoppe, Universität Magdeburg, 1994
3. Vorlesung „CGR I 3D Objekt-Repräsentation“, Michael Haller, Fachhochschule Hagenberg, 2004
4. Glossar vom Verband der Automobilindustrie
(http://www.vda.de/de/vda/intern/organisation/abteilungen/downloads/log_vda_4955_anl_1_de.pdf)
5. Vorlesung „Modellierung von Festkörpern“, Bernhard Jung und Ipke Wachsmuth, Universität Bielefeld
6. „Adhoc-Programmierung eines Volumenmodellierers im Bauwesen“, André Pichler, Universität Weimar, 2002
7. Vorlesung „Produktmodelle“, Lehrstuhl für Maschinenbauinformatik, Universität Magdeburg, 2001
8. Vorlesung „Modellierung“, Heidi Weber, Fachhochschule Vorarlberg, 2003/2004
9. Real-time Dynamic Deformable Meshes for Volumetric Segmentation and Visualisation, Bowden, R. Mitchell, T. A. Sahardi, M. , Vision and VR Group, Dept M & ES Brunel University Uxbridge, Middlesex UB8 3PH , 1997
10. Spin-Images: A Representation for 3-D Surface Matching, Andrew Edie Johnson, Carnegie Mellon University Pittsburgh, Pennsylvania 15213, 1997

11. Grafische Datenverarbeitung und Künstliche Intelligenz, Michaela Schmalzl und Martina Kassegger, Universität Wien, 2003