

Histograms of Oriented Gradients for 3D Object Retrieval

Philipp Lambracht

Abstract—Die 3D Objekterkennung ist ein wichtiges Themengebiet der Mobilien Systeme und der autonomen mobilen Robotik geworden. Ein populärer Ansatz um die Ähnlichkeit zwischen 3D Objekten zu bestimmen, sind globale Deskriptoren. Im Zuge meiner Ausarbeitung für das Proseminar „Mobile Systems Engineering“ habe ich den wissenschaftlichen Artikel „Histograms of Oriented Gradients for 3d Object Retrieval“ von Maximilian Scherer, Micheal Walter und Tobias Schreck ausgewählt und werde den darin beschriebenen und entwickelten Deskriptor, im folgenden HOG3D genannt, genauer vorstellen.

I. EINLEITUNG

Zum Zeitpunkt der Veröffentlichung des von mir ausgewählten Wissenschaftlichen Artikel wurden viele unterschiedliche Methoden zur 3D Objekterkennung vorgestellt. Einzelne Deskriptoren konnten sich bisher nicht als überlegen herausstellen. Es hat sich vielmehr etabliert jeweilige Deskriptoren geschickt zu kombinieren um deren Stärken zu nutzen und somit eine weitaus bessere Performance zu erzielen. Dementsprechend wird der in [6] vorgestellte HOG3D mit hoch-dimensionalen Merkmal Vektoren verglichen und eine Kombination mit diesem in einem Experiment versucht. Hauptmotivation für der HOG3D waren unter anderem bereits erfolgreiche Anpassungen von 2D Bild Analyse Methoden auf 3D Objekterkennung. Es wurde sich für die Anpassung des bereits erfolgreichen HOG aus [2] entschieden.

A. Grundbegriffe

Im folgenden werde ich ein paar wichtige Grundbegriffe für diese Ausarbeitung erläutern.

1) *Globaler und partieller Ansatz*: Bei der 3D Objekterkennung gibt es zwei verschiedene Ansätze. Der globale Ansatz betrachtet jeweils die komplette Form des 3D Modells und es werden nach Ähnlichkeiten gesucht, während der partielle Ansatz nach lokalen Ähnlichkeiten sucht. Hierbei ist zu beachten, dass es bisher keine absolute Lösung des Ähnlichkeitsproblems existiert, werde für den globalen noch für den partiellen Ansatz. Dementsprechend haben entsprechende Lösungsversuche einen heuristischen Natur. [6].

2) *Histogramm*: Histogramme dienen der in der Statistik und Bildverarbeitung dazu Häufigkeiten bestimmter Merkmale visuell darzustellen. Ein einfaches Beispiel aus der Bildverarbeitung wäre ein Histogramm eines Graustufenbildes mit den jeweils darin vorkommenden Grauwerten.

Tabelle II
GRAUWERTBILD ALS MATRIZE

a_{00}	a_{01}	a_{02}	a_{03}
a_{10}	100	50	235
a_{20}	73	42	150
a_{30}	30	125	0

Tabelle I
GRAUWERT HISTOGRAMM

Grauwert	Anzahl
150	30
20	5
...	...
255	10

Eine Detail reichere Einführung im Bezug auf Bildverarbeitung ist in [5] zu finden.

3) *Gerichtete Gradienten*: Gerichtete Gradienten werden wie z.B. in [2] äußerst erfolgreich zur Merkmaldetektion für 2D Bilder eingesetzt. Die Verwendung dieses Begriffs kann in [6] und dementsprechend in dieser Ausarbeitung vom mathematischen Begriff abweichen.

Um Gerichtete Gradienten zu berechnen, benötigt man Gradientenoperatoren. Hiermit sind Lineare Filter aus der Bildverarbeitung gemeint. In der Einführungslektüre [5] versteht man Filter als Funktionen welche auf Bilder, als Matrizen darstellbar, angewendet werden. Gradientenoperatoren sind gemäß der Definition über differenzierbare Funktionen, eine entsprechende Approximation mit denen man z.B. 2D Bilder „ableiten“ kann. Die Filtermaske $\begin{Bmatrix} -1 & 0 & 1 \end{Bmatrix}$ bewirkt z.B. die 1. Ableitung. Dieser Filter kann z.B. für ein 2D Bild dementsprechend in die X-Richtung und in die Y-Richtung angewendet werden.

Formel 1 zeigt ein Beispiel, wie ein Element aus Tabelle II abgeleitet wird. In diesem Fall in X-Richtung. An den Rändern muss jeweils eine Randbehandlung vorgenommen werden. Werte können z.B. gespiegelt werden.

$$a'_{22}x = -75 + 0 + 150 = 75 \quad (1)$$

Mit den Gradientenoperatoren lässt sich jeweils die Gradientenlänge bzw. -betrag und die Gradientenrichtung berechnen. Die Formeln 2 und 3, entnommen aus [5] zeigen jeweils die Berechnung für 2D Bilder. I_x bzw. I_y steht jeweils für die Ableitung in X- bzw. Y-Richtung. Mit dem Parameter p ist den entsprechenden Pixel gemeint.

$$G_l(p) = \sqrt{I_x^2(p) + I_y^2(p)} \quad (2)$$

$$G_r(p) = \arctan_2(-I_y(p), I_x(p)) \quad (3)$$

4) *3D Mesh*: 3D Meshs werden dazu verwendet um 3D Objekte digital zu speichern. Es werden Informationen über die Vertices (Punkte), Kanten, Flächen, Polygone sowie falls nötig Informationen über die Oberfläche (z.B. Farbe) gespeichert. In dem von mir Ausgewählten Artikel [6] werden Meshs aus schon bestehenden Performance-Tests genommen um die Leistungsfähigkeit des HOG3D zu messen.

II. HAUPTTEIL

Einige Referenzen sind [3], [1], [6], [2] und [4] und [5]

A. Meine erste Sektion

Tabelle ?? zeigt etwas.

B. Meine zweite Sektion

Abbildung 1 zeigt etwas.

Hier sollte ein Bild sein.

Abbildung 1. Inductance of oscillation winding on amorphous magnetic core versus DC bias magnetic field

III. DISKUSSION

REFERENCES

- [1] Sanjeev Arulampalam, Simon Maskell, Neil Gordon, and Tim Clapp. A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking. *IEEE Transactions on Signal Processing*, 50(2):174–188, 2002.
- [2] Navneet Dalal and Bill Triggs. Histograms of oriented gradients for human detection. In *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, volume 1, pages 886–893. IEEE, 2005.
- [3] Richard Hartley and Andrew Zisserman. *Multiple View Geometry in Computer Vision (2nd Edition)*. Cambridge University Press, Cambridge, UK, 2003.
- [4] Ralf Kaestner, Sebastian Thrun, Michael Montemerlo, and Matt Whalley. A non-rigid approach to scan alignment and change detection using range sensor data. In *Proceedings of the Symposium on Field and Service Robotics*, pages 179–194, Port Douglas, Australia, July 2005.
- [5] Lutz Priese. *Computer Vision*. Springer Vieweg, 2015.
- [6] Maximilian Scherer, Michael Walter, and Tobias Schreck. Histograms of oriented gradients for 3d object retrieval. 2010.