

appeared in
10. Workshop Farbbildverarbeitung 2004, Koblenz, Online-Proceedings
<http://www.uni-koblenz.de/icv/fws2004/>

Ein Entwurfswerkzeug für Farbklassifikatoren in Echtzeitanwendungen

Klaus Arbter und Daniel Kish

Deutsches Zentrum für Luft- und Raumfahrt e.V.
Institut für Robotik und Mechatronik Oberpfaffenhofen
D-82230 Weßling
klaus.arbter@dlr.de
www.robotic.dlr.de

Zusammenfassung. Farbsegmentierung in Echtzeitanwendungen konnte bis vor wenigen Jahren nur unter Verwendung teurer Hardware für Kameras, Framegrabber und Spezialcomputer realisiert werden. Heute ist leistungsfähige Hardware sehr viel billiger erhältlich, und die Leistung eines modernen Personal Computers ist in vielen Fällen ausreichend. Eine effiziente Software-Implementierung besteht aus einer Color-Look-Up-Tabelle (CLUT) gefolgt von einem morphologischen Filter. Dabei ist die Belegung der CLUT anwendungsspezifisch. Die richtige Belegung zu finden, kann sehr mühsam sein. In dieser Arbeit wird ein interaktives Entwurfswerkzeug vorgestellt, das über eine grafische Benutzeroberfläche bedient wird und es bequem ermöglicht, aus mehreren Bildern eine CLUT abzuleiten. Die Verwendung von Farbklassifikatoren in Echtzeitanwendungen wird an Beispielen aus der Robotik gezeigt.

1 Einleitung

Ein Zugang zur Lösung von Problemen der schritthaltenden Bildverarbeitung in Echtzeitanwendungen besteht in einer möglichst raschen Datenreduktion. Der erste Schritt ist dabei oft die Umwandlung des ankommenden Bildes in ein Kantenbild oder Binärbild, wobei unterstellt wird, dass Bildkanten zu Objektkanten und Bildregionen zu Objektflächen korrespondieren. Seit den frühen Tagen des maschinellen Sehens werden Grauwertbilder auf diese Weise vorverarbeitet. Der nahe liegenden Idee, Farbbilder entsprechend zu behandeln, standen bis vor kurzem sehr hohe Implementierungskosten für Farbkameras und für Spezialhardware zur Bildvorverarbeitung im Wege. Die rasante technologische Entwicklung der letzten Jahre hat dazu geführt, dass heute Farbkameras erhältlich sind, die weniger kosten als eine Tankfüllung Benzin, und dass die Rechen- und Datentransferleistung moderner PCs ausreichen, um die Farbklassifikation im Videotakt auszuführen [2].

Eine bewährte Implementierung des Farbklassifikators besteht aus einer Look-Up-Tabelle (CLUT) gefolgt von einem binären Ortsfilter (closing) [6] und einem zeitlichen Tiefpassfilter mit binärem Ausgang [9]. Die CLUT weist jedem Pixel aufgrund seiner

Farbe eine von zwei Bedeutungen zu (Farbe bekannt/unbekannt), sodass als Zwischenergebnis ein Binärbild entsteht. Die Anpassung an eine konkrete Aufgabenstellung erfolgt durch die Belegung der CLUT, die auch Klassifikatorentwurf genannt wird. Das kann vollautomatisch, unter Verwendung von beispielhaftem Bildmaterial, geschehen. Dazu muss zuerst das Entwurfsziel durch die Angabe von Entwurfsparametern festgelegt werden, was den Nichtexperten vor ein schwieriges Problem stellen kann. Dann muss der Anpassungsprozess durch den Anwender initialisiert werden, z. B. indem er Regionen markiert, in welchen wieder zu erkennende Farben vorkommen. Und schließlich benötigt er noch die Möglichkeit, das Ergebnis des automatischen Entwurfs zu korrigieren.

Es erscheint daher attraktiv, auf den vollautomatischen Entwurf zu verzichten und dafür Werkzeuge einzusetzen, die den interaktiven Entwurf unterstützen. In der vorliegenden Arbeit wird das interaktive Entwurfswerkzeug CCDT (Color Classifier Design Tool, DLR) vorgestellt. Auch hier beschreibt der Anwender im Bildraum die Regionen, welche die Farben enthalten, die wieder erkannt werden sollen. Die Klassengrenzen im Farbraum kann er weitgehend frei festlegen, aber auch an Schwellwerte in Farbhistogrammen koppeln. Zu jeder Zeit kann er seinen Entwurf an Testbildern überprüfen. Als Ergebnis einer Sitzung erhält er eine Datei, welche die CLUT in einem dicht komprimierten Format enthält. Diese Daten kann er unmittelbar in seine Anwendung übertragen. Anwendungsbeispiele aus der Robotik schließen die Arbeit ab.

2 Das Entwurfswerkzeug CCDT

Jeder interaktive Entwurf erfolgt im Wechsel zwischen den Entwurfsphasen Analyse, Synthese und Verifikation. Dementprechend ist das Entwurfswerkzeug CCDT (Color Classifier Design Tool) konzipiert. Es stellt Hilfsmittel zur Farbbildanalyse, zur Klassifikatorsynthese und zur Anwendung der aktuellen Klassifikatorauslegung an Testmaterial zur Verfügung. Diese sind funktional voneinander unabhängig und können daher in beliebiger Folge angewandt werden.

Die Benutzeroberfläche Die Benutzeroberfläche (Abb. 1) ist in drei Abschnitte gegliedert. Links befinden sich die Darstellungen des Bildraums *Image Space*, in der Mitte die des XY-Farbraums *Color Space* und rechts die des Z-Grauertraums *Intensity Space* sowie ergänzende Bedienelemente. Dem Bild- und dem Farbraum sind unabhängige Menüs zugeordnet. Als XY-Farbraum wird der kartesische, diskrete Farbraum auf der Displayebene aufgefasst. In diesen eingebettet ist der HS-Farbraum, weil dieser mit der menschlichen Farbwahrnehmung gut korrespondiert und sich daher für die Orientierung des Benutzers im Farbraum gut eignet. Der Z-Grauertraum, ist die diskrete Repräsentation der V-Komponente des HSV-Farbraums [8].

Farbbildanalyse Die Trainings- und Testbilder können in beliebigen Verzeichnissen gespeichert sein. Sie sind einzeln über einen Verzeichnisbaum erreichbar (*Load Image*). Die unterstützten Bildformate sind PNG, XPM, JPG, JPEG, BMP, PPM und GIF. Zur Farbbildanalyse werden die Bilder der Reihe nach vom Benutzer geladen und bearbeitet. Das Ergebnis der Analyse sind ein Farb- und ein Grauerthistogramm, welche die

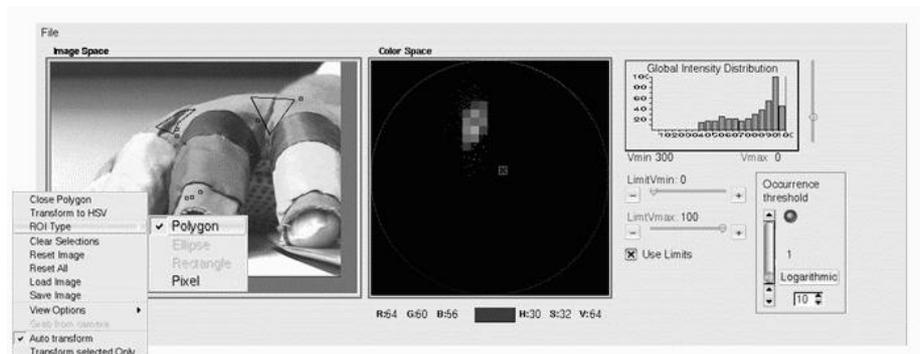


Abb. 1. Die CCDT-Benutzeroberfläche bei einer Farbanalyse: Farbbild mit benutzerdefinierten Regionen, Punkten und Menüs des Bildraums (links), Farbbelegung und überlagertes, begrenztes Farbhistogramm (Mitte), kumulatives Grauwert histogramm (rechts)

Grundlage für den nächsten Schritt, die Klassifikatorsynthese bilden. Nach dem Laden eines Bildes (*Load Image*) werden Bildpunkte (*ROI Type: Pixel*) oder Regionen (*ROI Type: Polygon*) oder das ganze Bild (*Click auf Bild*) markiert. Die anschließende Aufforderung *Transform to HSV* bewirkt, dass alle erfassten Pixel des Farbbildes in den XYZ-Farbraum transformiert und die Ergebnisse in ein XY-Farbhistogramm und ein Z-Grauwert histogramm eingetragen werden. Im Modus *Auto transform on* erfolgt die Transformation automatisch nach dem Abschluss jeder Markierung. Die Ergebnisse der Transformation können im Farbbereich auf verschiedene Weise dargestellt werden. In der Grundeinstellung werden alle ermittelten Bildfarben im XY-Farbraum in der zuletzt gefundenen RGB-Originalfarbe markiert. Diese Einstellung vermittelt eine schnelle Übersicht über die Belegung des Farbraums. Eine detailliertere Darstellung für die nachfolgende Klassifikatorauslegung bietet das Farbhistogramm (*View Options: Histogram Options: Histogram on*). Dieses zeigt in Falschfarbendarstellung die relative Farbhäufigkeit im XY-Farbraum, jedoch nur für die Werte, welche über dem einstellbaren Schwellwert *Occurrence threshold* liegen.

Klassifikatorsynthese und -verifikation Die Synthese des Klassifikators erfolgt durch die unabhängige Festlegung von Grenzen (*LimitVmin*, *LimitVmax*) im Z-Grauwertraum und von Farbwerten im XY-Farbraum. Die Farbwerte können auf unterschiedliche Weise spezifiziert werden:

- (a) alle Farbwerte, die vom begrenzten Farbhistogramm bedeckt werden
- (b) alle Farbwerte die von benutzerdefinierten Polygonen eingeschlossen werden
- (c) alle Farbwerte, die von benutzerdefinierten Punkten getroffen werden

In einfachen Fällen reicht die automatische Auslegung nach (a) aus (*Filter Options: Auto Select on*). Sie kann aber auch ausgeschaltet (*Filter Options: Auto Select off*) und eine manuelle Auslegung nach (b) und (c) durchgeführt werden. Dabei bietet die bloße Anzeige des begrenzten Farbhistogramms (*View Options: Histogram Options: Histogram on*) eine wertvolle Orientierungshilfe. Die benutzerdefinierten Elemente (Polygon

und Punkt) sind wahlweise konstruktiv (*Add to CLUT*) oder destruktiv (*Remove from CLUT*). Damit sind im Laufe der Auslegung immer wieder Korrekturen möglich. In jeder Phase des Entwurfs kann die Wirkung der aktuellen Auslegung auf das jeweils geladene Bild überprüft werden (*Apply Filter*). In der Grundeinstellung werden die erkannten Pixel einfarbig markiert, im *Inverse Filter* Modus alle anderen, nur nicht die erkannten. Ein Beispiel zeigt Abbildung 3.

Anmerkungen zur Implementierung Die interne Bilddarstellung erfolgt grundsätzlich im RGB-3x8Bit-Format. Bilder von geringerer Auflösung (z. B. 565-Format) werden automatisch expandiert. Damit ist eine konsistente, formatunabhängige Verarbeitung gewährleistet.

Ein Problem, das besondere Beachtung verdient, sind die Abbildungsfehler in diskreten Räumen. Während die kontinuierliche Abbildung $RGB \rightarrow HSV$ umkehrbar eindeutig ist, ist es die diskrete Abbildung $RGB \rightarrow XYZ$ nicht. Zu einem Punkt \mathcal{P} im XYZ-Raum gibt es im Allgemeinen eine Region \mathcal{R} im RGB-Raum, deren Abbildung auf \mathcal{P} führt. Das heißt, dass bei der Auswahl eines Punktes im XY-Raum, und eines Intervalls im Z-Raum alle direkten Nachbarn einer Geraden im diskreten RGB-Raum, sowie alle entfernteren, potenziellen Nachbarn gefunden werden müssen. Die vollständige Suche scheidet aus Effizienzgründen aus. Die direkten Nachbarn der Gerade zu finden, ist nicht schwer. Schwieriger ist es, alle anderen potenziellen Nachbarn, aber nur diese, zu lokalisieren. In der aktuellen CCDT-Version ist dieses Problem noch nicht exakt gelöst, sondern hier werden alle Punkte in einer ausreichend großen Umgebung der direkten Nachbarn zu Nachbarn ernannt. Der daraus resultierende Klassifikator lässt geringfügig mehr Farben zu, als er es exakterweise sollte.

3 Anwendungen

Multisensorielles Robot Servoing Bei Greif- oder Dockingvorgängen von Robotern ist die Annäherung an eine Nominalposition relativ zum Objekt eine wichtige Grundaufgabe, die bei Objekten mit variabler Lage nur durch den Einsatz von Sensoren gelöst werden kann. In [3] wird ein analytischer Ansatz und in [10] ein künstliches neuronales Netz zur Verarbeitung redundanter Information aus einem Stereo-Kamerapaar und vier Laser-Entfernungsmessern vorgestellt. Beide wurden unter Verwendung des multisensoriellen DLR-Robotergrifiers (Abb. 4a) getestet. Als Bezugspunkte für die Bildverarbeitung dienten vier Marken auf dem Objekt, die zuerst als schwarze Punkte auf hellem Grund ausgeführt waren. In der komplexen Experimentalumgebung (Abb. 4b) mit wechselnden Schrägansichten bereitete deren Erkennung gelegentlich Schwierigkeiten, die mit der Verwendung von Farbmarken vollständig behoben werden konnten. Die Abbildung 4c zeigt ein Klassifikationsergebnis der CLUT, in welchem die Restfehler durch eine nachgeschaltete Filterung leicht unterdrückt werden können.

Kalibrierung eines Datenhandschuhs Ein sehr aktuelles Thema in der Roboterforschung ist die Entwicklung anthropomorpher Roboterhände und deren Kontrolle zur Bewältigung komplexer Manipulationsaufgaben. Fischer et. al. [4] beschreiben einen Ansatz zur Abbildung der nichtlinearen und verkoppelten Messwerte eines

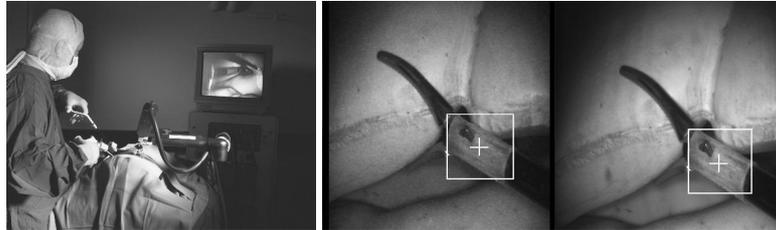


Abb. 2. Solochirurgische Szene mit Assistenzroboter (links) und Stereo-Bildpaar mit umschreibenden Rechtecken der segmentierten Farbmarke (rechts)

Datenhandschuhs auf die Kinematik der DLR-Hand (Abb. 5a) mittels eines künstlichen neuronalen Netzwerks. Zum Training des Netzwerks müssen viele Konfigurationen der Fingerspitzen des Datenhandschuhs zusammen mit den jeweiligen Messwerten erfasst werden. Diese Aufgabe wird mithilfe farbiger Stecknadeln (Abb. 5b) an den Fingerspitzen, deren räumliche Positionen durch Stereobildauswertung ermittelt werden, gelöst. In Abbildung 5c ist die Raumbedeckung durch die Fingerkuppen des Datenhandschuhs dargestellt.

Roboterassistent für die laparoskopische Chirurgie Seit einigen Jahren werden minimalinvasive Operationstechniken entwickelt und zunehmend klinisch angewandt. Während der Chirurg beim offenen Eingriff den Situs unmittelbar sieht, blickt er beim minimalinvasiven Eingriff auf einen Monitor, auf dem das Bild eines Laparoscops dargestellt ist. Dazu muss ein Kollege (Assistent) das Laparoskop halten und kontinuierlich am Situs ausrichten. Diese Tätigkeit ist anstrengend und mit vielen Problemen behaftet. Deshalb wurde ein automatisches Laparoskopführungssystem entwickelt [1, 9], bei welchem ein Roboter ein Stereolaparoskop hält (Abb. 2) und selbstständig auf eine Farbmarke (Abb. 6a) an der Instrumentenspitze ausrichtet (Abb. 6b). Der Klassifikator-entwurf erfolgte unter den folgenden, erschwerenden Bedingungen:

- hoher Störpegel (Rauschen) in den Laparoskopbildern
- Material der Marke muss biokompatibel sein
- schwach gesättigtes Farbbild der Marke (Abb. 6f) und
- teilweise überlappende Farbcluster (Abb. 6e und 6f)

Erleichternd wirkte sich aus, dass auf eine Auswertung der Form der Marke in der Vorverarbeitung verzichtet wurde, und daher auch Klassifikationsfehler auf dem Bild der Marke großzügig toleriert werden konnten. Zur Nachbearbeitung (Abb. 6g und 6h) wurden ein Ortsfilter (Closing), ein Zeitfilter (Tiefpass) und eine einfache Blob-Analyse verwendet. Das System wurde ursprünglich unter Verwendung sehr teurer Bildverarbeitungshardware (Datacube, MaxVideo) realisiert und klinisch getestet [7]. Zurzeit ist eine PC-basierte Realisierung in der Entwicklung.

Visuelle Erfassung der Herzbewegung Herzoperationen erfolgen üblicherweise am stillgelegten Herzen, wobei der Kreislauf durch eine Herz-Lungen-Maschine

aufrechterhalten wird. Um die Nachteile dieser Maschine zu vermeiden, werden neuerdings Methoden für den robotergestützten Eingriff am schlagenden Herzen untersucht. Eine zentrale Rolle spielt dabei die Erfassung der Restbewegung des mechanisch stabilisierten Herzens, um eine abgestimmte Bildstabilisierung für den Chirurgen und Bewegungskompensation des Roboters zu realisieren. Zur visuellen Bewegungsschätzung wurden von Gröger et. al. lokaler Grauwertstrukturen als natürliche Landmarken und die Eignung farbiger Titan-Clips als künstliche Landmarken untersucht [5]. In den Abbildungen 7a bis 7f sind die Szene und eine Klassifikatorauslegung dargestellt.

4 Zusammenfassung und Ausblick

Es wurde ein interaktives Entwurfswerkzeug für Farbklassifikatoren vorgestellt, mit welchem Nichtexperten einfachere Auslegungen bequem und schnell vornehmen können, welches es aber auch Experten erlaubt, ihre Expertise einzubringen. Die Beispiele aus der Robotik zeigen ein breites Einsatzspektrum der Farbklassifikation in Echtzeitanwendung und damit auch den potenziellen Nutzen des vorgestellten Werkzeugs im Labor und in Projekten. In der Folgezeit wird das Werkzeug weiterentwickelt, wobei insbesondere die Erfahrung aus der Praxis in die Entwicklung einfließen soll.

Literatur

1. K. Arbter and G.Q. Wei. Verfahren zum Nachführen eines Stereo-Laparoskops in der minimalinvasiven Chirurgie. Deutsche Patent Nr. 195 29 950, 14. Nov., 1996.
2. James Bruce, Tucker Balch, and Manuela Veloso. Fast and inexpensive color image segmentation for interactive robots. In *Proc. 2000 IEEE/RSJ International Conference on Intelligent Robots and Systems*, volume 2, pages 2061–2065. IEEE, 2000.
3. B. Brunner, K. Arbter, and G. Hirzinger. Task directed programming of sensor based robots. In *Proc. IEEE Int. Conf. on Intelligent Robots and Systems (IROS'94), München, Sept. 12 - 16, 1994*.
4. M. Fischer, P. van der Smagt, and G. Hirzinger. Learning techniques in a dataglove based telemanipulation system for the DLR hand. In *Proc. Int. Conf. on Robotics and Automation, ICRA '98, Leuven, Belgium, 1998*.
5. Martin Gröger, Bernhard Kübler, and Gerd Hirzinger. Selection of qualified colour markers for heart surface tracking. In *VDE Proceedings of the 38th Annual Congress of the DGBMT, Ilmenau, Germany, Sept. 2004*.
6. R. Massen, T. Regele, and P. Böttcher. Colour and shape classification with competing paradigms: Neural networks versus trainable table classifiers. In *ECO 3 Int. Congress Optic. Science, Den Haag, 1990*.
7. K. Omote, A. Ungeheuer, H. Feussner, K. Arbter, G.Q. Wei, J. R. Siewert, and G. Hirzinger. Self-guided robotic camera control for laparoscopic surgery compared with human camera control. *American Journal of Surgery*, 177:321–324, April 1999.
8. Alvy Ray Smith. Color gamut transform pairs. In *Proceedings of the 5th annual conference on computer graphics and interactive techniques, SIGGRAPH '78*, pages 12–19. ACM Press, Aug. 1978.
9. G.-Q. Wei, K. Arbter, and G. Hirzinger. Real-time visual servoing for laparoscopic surgery. *IEEE Engineering in Medicine and Biology*, 16(1), January/February 1997.
10. G.Q. Wei and G. Hirzinger. Learning motion from images. In *Proc. The 11th Int. Conf. Pattern Recognition*, pages 189–192, 1992.

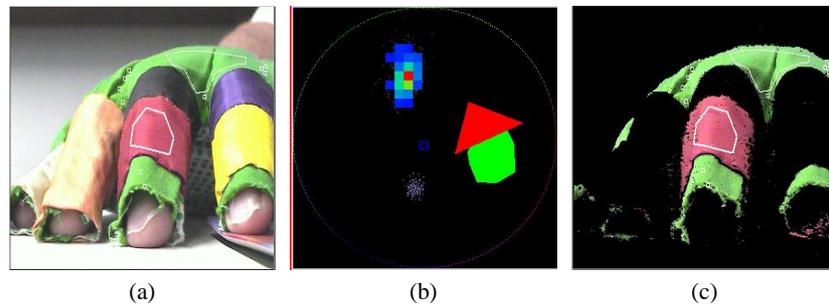


Abb. 3. Beispiel für den Klassifikatorentwurf: (a) Farbbild mit eingetragenen Polygonen in den grünen, roten und violetten Bildregionen. (b) Beschränktes Farbhistogramm beim Farbton grün, konstruktive (grüne) Fläche beim Farbton rot, destruktive (rote) Fläche beim Farbton rötlich gelb und Cluster beim Farbton violett. (c) Die roten und grünen Farben werden erkannt (im Modus: *Auto Select On*), die rötlich gelbe Farbe wird explizit unterdrückt (Finger links), und die violette wird ignoriert.

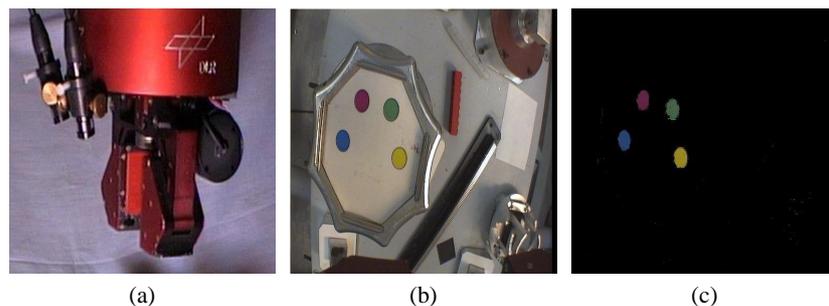


Abb. 4. Bildauswertung für den multisensoriell gestützten Greifvorgang eines Roboters: (a) Multisensorieller DLR-Greifer (Stereo-Kamerapaar, vier Laser-Entfernungsmesser in den Greiferbacken), (b) Farbmarkiertes Objekt (ORU: Orbital Removable Unit) mit rötlichen Objekten im Hintergrund der ROTEX-Arbeitszelle, (c) segmentierte Farbmarken des ORU (ohne Filterung)

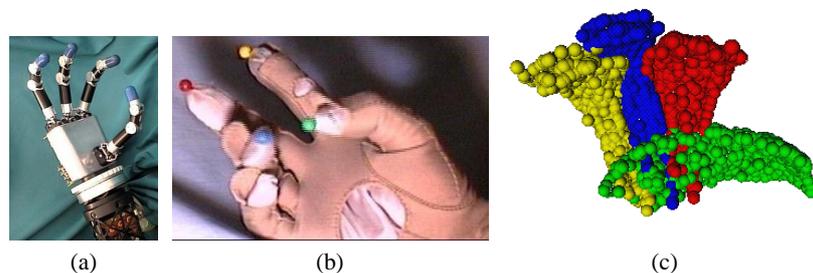


Abb. 5. Farbsegmentierung für die Kalibrierung eines Datenhandschuhs zur Telemanipulation einer Roboterhand: (a) DLR-Hand (erstes Experimentalmodell), (b) Datenhandschuh mit Farbmarken an den Fingerspitzen, (c) aus Stereo-Bildfolge experimentell ermittelter Arbeitsraum der Fingerkuppen während eines Versuchs

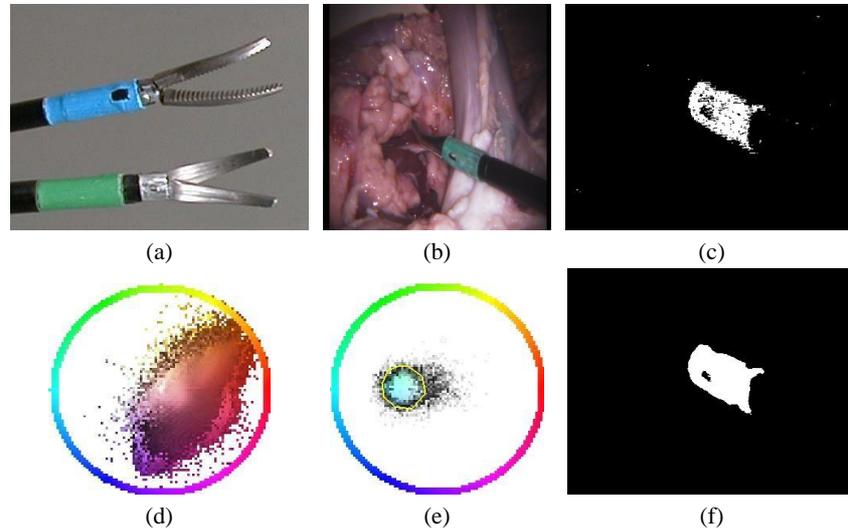


Abb. 6. Farbsegmentierung für die bildgestützte, automatische Laproskopführung durch einen Chirurgieroboter: (a) farbcodierte Instrumente, (b) Situs mit Instrument und Farbmarke, (c) Klassifikationsergebnis ungefiltert, (d) HS-Histogramm des Hintergrunds, (e) HS-Histogramm der Farbmarke, (f) Klassifikationsergebnis gefiltert

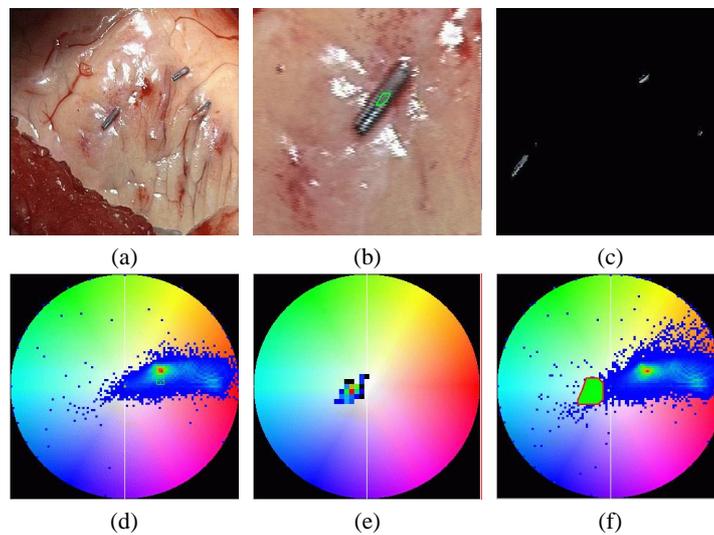


Abb. 7. Segmentierung farbiger Clips auf der Herzoberfläche für die Bewegungskompensation bei chirurgischen Eingriffen am schlagenden Herzen: (a) Herzoberfläche mit farbigen Clips, (b) grafisch markierte Region auf einem Clip, (c) segmentierte Clips (vergrößerter Ausschnitt, ohne Filterung), (d) Farbhistogramm der Herzoberfläche, einschließlich Clips, (e) Farbhistogramm der markierten Region in (b), (f) Klassifikator Auslegung: rot umrandete grüne Fläche