

Analyse und Entwicklung optischer Trackingverfahren für das dreidimensionale Skizzieren in der virtuellen Produktentwicklung

Patrick Tobias Fischer

PatrickTobias.Fischer-projekt@ipk.fraunhofer.de
mail@fischerkinder.de

Fachhochschule Köln
University of Applied Sciences Cologne
Department Computer Sciences and Engineering

Fraunhofer-Institute for Production Systems
and Design Technology IPK
Department Virtual Product Creation

Abstract. Sketching in three-dimensional immersive environments is seen as a new way for externalizing mental models and creating ground-breaking ideas for new products. While the sketching process in the 3d-space seems to generate some benefits like being able to move throughout the sketch and interact with it or sketching objects in 1:1 scale, the technology making 3d-sketching possible may limit the designer's skills in a variety of ways.

The purpose of the thesis is to analyze the process of 3d-sketching in a holistic way by analyzing theory and real-life to find a way to lower the technical complexity that limits inherently the designer's skills. The solution proposed here especially fits the properties of a CAVE since it is using only one camera for determining translation and orientation in 3d-space, by using the POSIT-Algorithm.

1. Einleitung und Zielsetzung

Bereits in den frühen Phasen der Produktentwicklung werden 80% der Kosten eines Produktes festgelegt. Umso wichtiger ist es in der Entwurfsphase, einen möglichen Lösungsraum für eine Produktidee weitestgehend zu erschöpfen. Skizzen sind dabei eine gute Methode schnell und effizient Lösungen zu entwickeln. Einerseits helfen sie mentale Modelle zu externalisieren; andererseits ermöglichen sie dem Designer, Konstrukteur oder Architekten, das Problem mit sich selbst zu diskutieren (Tversky 2003).

Bestehende CAS-Systemen (*Computer Aided Styling*), welche im virtuellen Entwicklungsprozess bereits langjährige Anwendung finden, erlauben aufgrund ihrer komplexen Benutzerschnittstelle nur selten eine spontane Umsetzung einer Idee. Der Stift als Interaktionsgerät ist eben weitaus intuitiver zu benutzen und damit kreativitätsfördernder, als Maus, Tastatur und GUI. Auch fällt die Beurteilung des Designs am Bildschirm bei CAS erstellten Skizzen schwieriger als beispielsweise bei einem klassischen Tonmodell.

Dennoch bieten CAS-Systeme viele Vorteile, die eine herkömmliche Papierskizze aufgrund ihrer Physikalität nicht bietet, wie z.B.:

- Modellerstellung basierend auf Konstruktionsdaten
- Dreidimensionalität und verschiedene Ansichten
- Variantenvielfalt
- Zurückgehen auf ältere Varianten

Die prototypische CAVE-Anwendung *SketchApp* versucht dieses Vorteile für die Skizze einträglich zu machen und lässt es dem Designer ermöglichen mittels magnetischem Tracking Striche im dreidimensionalen Raum mittels aktiver, stereoskopischer Projektion innerhalb einer CAVE entstehen zu lassen. Wie sich der Gesamtprozess darstellt, wie eine Verbesserung des Trackings möglich wäre, und welche Probleme (besonders in technischer Hinsicht) dabei auftreten, war Teil der Masterarbeit deren Ergebnisse in verkürzter Weise nachfolgend dargestellt werden.

2. Grundlagen des kreativen Entwurfsprozesses

Maßgebliches Ziel in der Entwurfsphase ist das Treffen einer Designentscheidung aus einer Vielzahl möglicher Lösungen. Dabei ist es erwünscht grade am Anfang der Produktentwicklung möglichst viele Lösungen in Betracht zu ziehen und den Lösungsraum weitestgehend zu erschöpfen, bevor kostspieligere Produktionsmaßnahmen angegangen werden. Das Lösen eines bestimmten Problems verläuft dabei in verschiedenen Stufen ab.

Ein Vergleich mehrerer Modelle zum Thema Problemlösen hat ergeben, dass *Problemdefinition* und *Beurteilung* wesentliche Schritte zur Entwicklung möglicher Lösungen sind. Gerade hier ergeben sich im Designprozess meist verschiedene skizzenhafte Darstellungen des Problems, die man als Varianten bezeichnet. Abbildung 2 und Abbildung 3 zeigen solche *Varianten*, die während eines Entwurfsprozesses entstanden sind. Der Designer zeichnet das Objekt mehrfach, aus verschiedenen Perspektiven, um eine Vorstellung darüber zu Bekommen, wie ein solches Produkt im Realen zu konstruieren wäre.

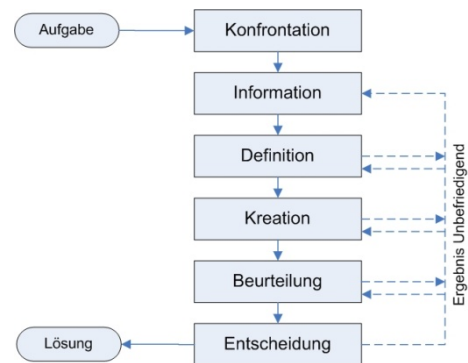


Abbildung 1 Der Konstruktionsprozess nach Pahl & Beitz

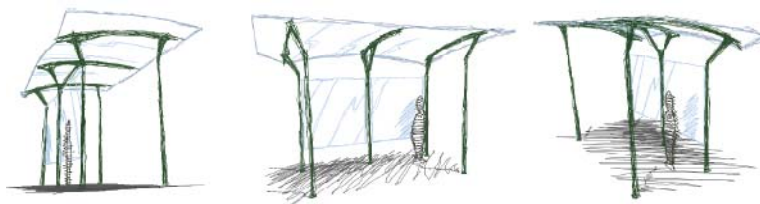


Abbildung 2 perspektivische Varianten



Abbildung 3 bauliche Varianten

Weiterhin ist in dieser Phase zu beobachten, dass die Designlösung gerne im *Kontext* gesehen wird. Abbildung 2 beispielsweise beinhaltet neben dem Designproblem „Bushaltestelle“ auch den wartenden Bürger an der Bushaltestelle selbst. Mit diesem Trick setzt der Designer seine Zeichnung in Größenrelation zu bekannten Objekten.

Untersuchungen und Experimente mit *SketchApp* haben zudem gezeigt, dass oft auch mit den gezeichneten 3d-Objekten „testweise“ interagiert wird. Der *Beurteilungsprozess*, der wie in Abbildung 1 vor der Entscheidung steht, gewinnt also an Qualität, da das 3d-gezeichnete Objekt in *SketchApp* eine 1:1 Abbildung des Produktes sein kann und das System eine *Interaktion* mit der Skizze selbst ermöglicht.

3. Human-Computer-Interaction

Während der Arbeit mit SketchApp und im Zuge der aktuellen wissenschaftlichen Diskussion im Bereich Tangible-User-Interfaces (TUI), entwickelte sich die Frage, ob der Stift von SketchApp als TUI zu bezeichnen ist, oder als herkömmliches Interface anzusehen ist. Auf Basis eines Modells von Hiroshi Ishii (Hiroshi Ishii 1997) wurde im Rahmen der Masterarbeit das folgende Modell entwickelt.

Kern des Modells sind zwei wesentliche Feststellungen, wenn es um die Kommunikation bzw. Interaktion zwischen Mensch und Computer geht:

1. Ausgabegerät (Display) und Eingabegerät (Interface) sind bei herkömmlichen Interaktionsparadigmen zwischen Mensch und Computer technologisch getrennt. Meist gibt es einen Ausgabekanal, realisiert über ein Monitor und einen Eingabekanal, bzw. -gerät, realisiert über Tastatur und Maus. TUIs vermischen Display und Interface miteinander und vereinen Eingabe- und Ausgabekanal miteinander, so dass für den Benutzer kein Unterschied zwischen Beidem besteht. Abbildung 5 und Abbildung 6 verdeutlichen dies grafisch.
2. Durch die Vereinigung von Display und Interface ergibt sich eine engere Kopplung zwischen dem Menschen mit seinem mentalen Raum und dem Computer mit seinem virtuellen Raum. Im Falle von SketchApp geschieht dies durch den hybriden Zeichenstift, bei dem ein Teil (Griff) tatsächliches physikalisches Eingabegerät ist aber die virtuelle Spitze von der CAVE (Display) angezeigt wird. Für den Benutzer gehört die Spitze trotzdem zum gefühlten Eingabegerät obwohl es technisch eine Ausgabe ist. Somit rutscht der Benutzer in seinem physikalischen Raum näher in die Virtualität. Und da das Skizzieren einem externalisieren von mentalen Modellen dient, ist der gefühlte Übergang zwischen mentalem Raum und virtuellem Raum wie es Abbildung 6 zeigt ebenso verringert worden.

Beobachtungsraum und Handlungsraum fallen zusammen und können so evtl. neue kreative Potentiale ermöglichen, da wie im Falle der Skizze die Selbstkommunikation unterstützt wird, aber zusätzlich eine weitere räumliche Dimension zur Verfügung steht.

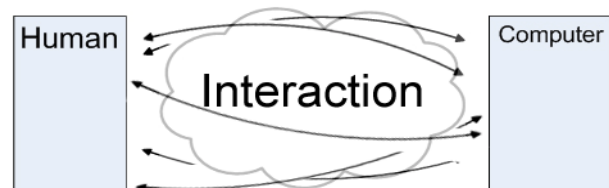


Abbildung 4 HCI Modell

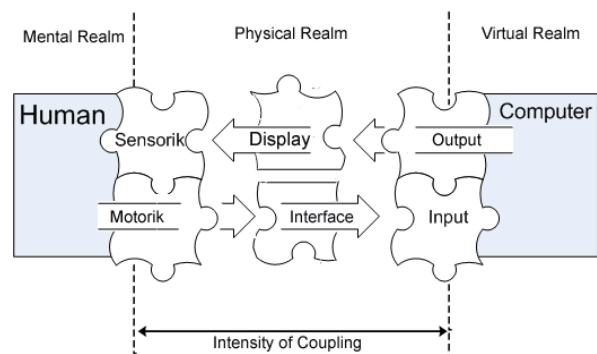


Abbildung 5 Transitionen von Informationen zwischen Mensch und Computer

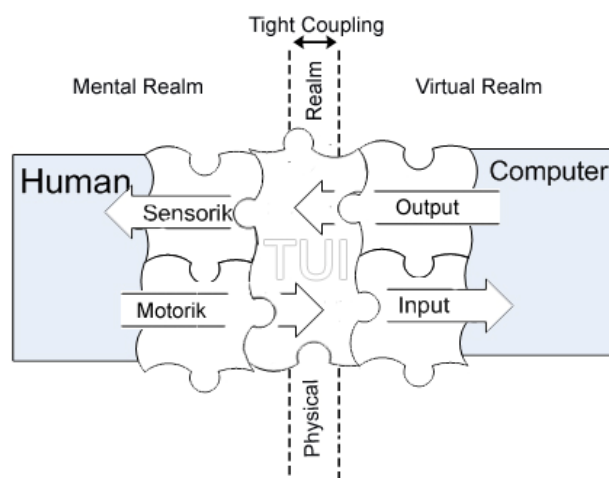


Abbildung 6 Engere Kopplung der Räume durch TUI

4. Analyse der Skizzierbewegungen und technische Herausforderungen

Das Tracking der 3d-Skizzieranwendung SketchApp wird derzeit über ein magnetisches Trackingsystem realisiert, welches durch metallische Objekte maßgeblich beeinflusst wird und daher eine präzise Synchronisierung von realem Raum und virtuellen Raum verhindert. Wichtig ist dies für SketchApp, da eine enge Kopplung zwischen Mensch und Computer erreicht werden soll, denn Verzögerungen und geringe Auflösung bremsen den Selbstkommunikationsprozess und hindern den Designer daran gleichzeitig zu denken und zu skizzieren (Römer 2002).

Trotz der verzerrten Positionsdaten, konnten durch das des magnetischen Trackingsystem wichtige Erkenntnisse über die Skizzierbewegungen innerhalb der Cave gewonnen werden. So ist beispielsweise beim normalen Skizzieren eine Handgeschwindigkeit von ca. 45 bis 250 cm/s zu erwarten. Bei der zur Verfügung stehenden Hardware (Unibrain Fire-I 620c) ergibt sich somit ein möglicher Abstand von ca. 1,5 bis 8,5cm von einem Koordinaten-Sample zum anderen.

Um die Grenzen der möglichen Handgeschwindigkeiten beim Skizzieren zu ermitteln, wurden Tests mit einer Hochgeschwindigkeitskamera durchgeführt. Die hier auftretenden Geschwindigkeiten bei der Translation befanden sich zwischen ca. 316 bis 745 cm/s und in der Rotation zwischen 49 cm/s bis 482cm/s (Bahngeschwindigkeit).

Weiterhin wurden im Rahmen dieser Arbeit die Aufenthaltsbereiche der Skizzierenden untersucht, um evtl. Hinweise über ein mögliches Kamerasetup für das optische Tracking zu erhalten. Es stellte sich heraus, dass ein sinnvoller Kameraaufbau innerhalb der CAVE nur unter Umständen möglich ist. Da es sich bei der CAVE um eine fünf Seiten CAVE handelt, sollte so wenig zusätzliche Hardware innerhalb der CAVE angebracht werden, um die Anzeigeeigenschaften so wenig wie möglich einzuschränken. Unter anderem aus diesem Grund wurde sich beim Entwurf der Trackinganwendung Optitrack für ein monoskopisches Tracking entschieden. Auf diese Weise kann die Verdeckung der Projektionsfläche minimiert werden, da weniger Kameras innerhalb der Cave installiert werden müssen.

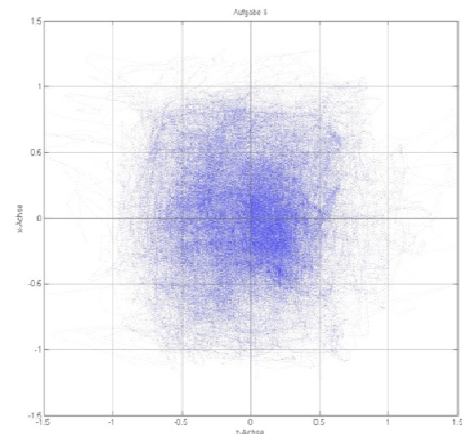


Abbildung 7 genutzter Interaktionsraum der CAVE

Wie aus Abbildung 7 hervor geht, sollten bei einem zukünftigen Kameraaufbau vor allem die Mitte der CAVE durch das Tracking unterstützt werden.

Ein theoretische Kameraaufbau wie er in Abbildung 8 gezeigt wird (45° Öffnungswinkel), erwies sich bei näherer Analyse als ungünstig, da der Skizzierende zu oft mit dem Rücken zur Kamera stehen und somit den Stift mit aufgesetztem Target verdecken würde. Zwar kann meist eine zweite Kamera den Stift evtl. trotzdem aufnehmen, aber durch einen geringen Öffnungswinkel sind auch aus diesem Sichtwinkel Einschränkungen (besonders in der Höhe) festzustellen.

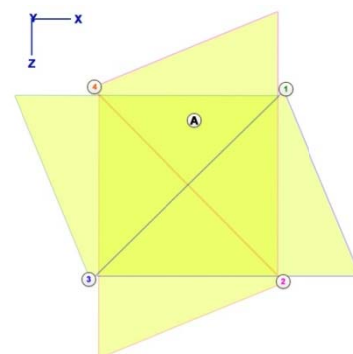


Abbildung 8 theoretischer Kameraaufbau

Derzeit nutzt das Optitrack ein Weitwinkel Kameraobjektiv mit einem Öffnungswinkel von 80° und einem IR-Hochpass Filter. Dies verbessert die unter Abbildung 8 dargestellte Situation, da aus einer Ecke der Cave heraus nahezu die gesamte CAVE überblickt werden kann. Der Nachteil der sich allerdings hieraus ergibt ist die notwendige Entzerrung des entstandenen, fischaugenartigen Kamerabildes. Die Funktion für das Entzerren eines 1024x768px Graubildes benötigt bei derzeitiger Hardwarekonfiguration in etwa 33ms.

5. Lösungsansätze

Während der Entwicklung von Optitrack, dem optischen Trackingsystem für SketchApp, wurden verschiedene Lösungsansätze prototypisch ausprobiert. Das Zielsystem wurde besonders in Hinblick auf die kritischen Anforderungen der CAVE hin konzipiert und immer wieder überdacht.

Eine erste Version von Optitrack versuchte das Target per *Color-Tracking* zu verfolgen. Die C-Bibliothek OpenCV stellt hierzu die Funktionen, bzw. Algorithmen Meanshift und Camshift zur Verfügung. Der Meanshift-Algorithmus war zwar schneller in seiner Durchführungszeit, konnte aber aufgrund seiner starren Suchfenster die einzelnen Kugeln des Multipunktmarkers nicht zuverlässig tracken. Der vermeintliche Vorteil eines dynamischen Suchfensters wie beim Camshift-Algorithmus erwies sich ebenfalls als unbrauchbar, da bei Entfernungen über einem Meter die Suchfenster zu klein wurden und schnellere Bewegungen aufgrund der kleinen Suchfenstergröße nicht mehr verfolgt werden konnten. Weiterhin wurde bei Testaufnahmen in der CAVE festgestellt, dass diese trotz der Projektionen relativ dunkel ist, und somit Farben schwer zuverlässig zu erkennen sind. Eine gewisse Kompensierung konnte hierbei der Einsatz von neon-farbenen Kugeln für den Multipunktmarker bieten.

Eine weitaus elegantere Methode, die Kugeln des Multipunktmarkers zu erkennen, wurde durch den Einsatz von Infrarotlicht bzw. Infrarotfilter erreicht. Auf diese Weise konnten auf analogem Weg alle irrelevanten Informationen im Bild entfernt werden. Es blieben lediglich die vier IR-Licht reflektierenden Kugeln im Bild zurück, welche mittels Thresholding und Erosion sowie Dilatation digital aufbereitet und zu einem Binärbild verarbeitet werden. Aus diesem Binärbild lassen sich schließlich sog. Blobs erkennen und deren Mittelpunkt bestimmen.

Die in Optitrack verwendete Methode lässt sich derzeit als *Blob-Detection* bezeichnen, da keine Informationen über die Identitäten der verschiedenen Kugeln zur Verfügung stehen. Es ist lediglich bekannt, dass in Position (x,y) ein weiser Punkt vorhanden ist. Um aus diesen Punkten Translation und Rotation des 4-Punkt-Markers zu errechnen, bedarf es dem von DeMenthon entwickelten SoftPOSIT-Algorithmus (DeMenthon 2003).

Da dieser Algorithmus bisher in keiner Programmier Bibliothek zu finden war, wurde eine eigene Implementierung in C unter Zuhilfenahme der OpenCV Bibliothek erstellt. In dieser Implementierung wurde klar, dass die Durchführungszeit mit 141ms sehr hoch ist und somit für eine Echtzeitanwendung nicht praktikabel. Es blieb die Verwendung des POSIT-Algorithmus (DeMenthon & Davis 1995), der allerdings im Gegensatz zum SoftPOSIT eine Korrespondenz zwischen Blobs in der Kamerabildebene und Kugeln im realen Raum benötigen.

Hier wurde mit einem kleinen Trick gearbeitet, der das Erkennen eines Referenzpunktes ermöglicht. Der 4-Punkt-Marker enthält eine Kugel, die 0,5cm größer ist als die restlichen. Auf diese Weise ist es möglich die Korrespondenz zuverlässig her zu stellen.

Abbildung 9 zeigt die Tracking-Pipeline für Optitrack. Derzeit wird ein Socket Server für Optitrack unter der Verwendung der Boost- und Asio-Bibliothek implementiert, um ermittelte Translation und Rotation des Zeichenstiftes dem CAVE-Rendercluster zu übergeben.

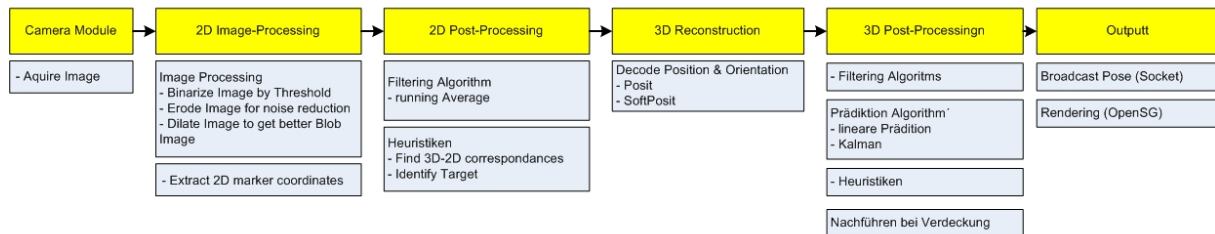


Abbildung 9 Tracking Pipeline

Ein Testbenchmark hat ergeben, dass die Bildaquirierung unter 1ms benötigt. 2D Image-Processing benötigt 53ms wobei davon 33ms für die Entzerrung des Bildes und 20ms für Thresholding, Erosion, Dilatation und Binarization benötigt werden.

Das 2D Post-Processing benötigt 30ms für die Extraktion der Blobs und Finden des größten Blobs. Die Durchführungszeit des Posit-Algorithmus liegt bei unter 1 ms.

6. Zusammenfassung und Ausblick

Um neue Interaktionsparadigmen schaffen zu können müssen ganzheitliche Überlegungen angestellt werden. Gerade im Bereich der HCI fällt es Technologen schwer benutzerzentriert zu denken und somit neue Potentiale durch digitale Technik zu schaffen. Ein ganzheitlicher Ansatz, wie ihn auch diese Arbeit bietet, ist in der Lage bessere Handlungsempfehlungen für weitere Forschungsvorhaben zu formulieren und Sinnhaftigkeit bestimmter Technologien zu bewerten.

SketchApp bietet das Potential solche neuen

Interaktionsparadigmen entstehen zu lassen. Der Hybride-

Zeichenwerkzeuge-Ansatz ist hierbei ein Anfang, wobei die Vorteile dieser noch genauer zu erforschen sind. Es gilt das Vorhandensein bzw. das Fehlen von Physikalität für diese neuen Interaktionsparadigmen bewusst einzusetzen.

Zur Weiterentwicklung von SketchApp und Optitrack lassen sich folgende Wege vorschlagen:

- Bessere Hardware ermöglicht eine engere Kopplung zwischen Mensch und Computer. Konkret betrifft dies vor allem die eingesetzte Firewire Kamera. Sie sollte durch eine Kamera mit höherer Bildrate und besserem Objektiv und Filter ausgetauscht werden.
- Der SoftPOSIT Ansatz ist weiter zu verfolgen, da Optimierungen im derzeitigen Code vermuten lassen, das Echtzeitfähigkeit möglich ist.
- Der Optitrack Server sollte vervollständigt werden, um ein von SketchApp autonomes System zu bekommen. Die Vernetzung wird beim Einsatz von mehreren Kameras ohnehin nötig sein.
- Eine Lösung um verschiedene Multipunktmarker mit einer Kamera zu erkennen ist zu finden.



Abbildung 10 Smear Effect

Literatur (Auswahl)

- 1 DeMenthon, D. F. (2003). *Matlab SoftPOSIT Demonstration Code*. Retrieved 04. March 2008, 2008, from www.cfar.umd.edu/~daniel/SoftPOSIT.txt.
- 2 DeMenthon, D. F. & L. S. Davis (1995). Model-Based Object Pose in 25 Lines of Code. *International Journal of Computer Vision* 15(1-2): 131-141.
- 3 Hiroshi Ishii, B. U. (1997). Tangible Bits: Towards Seamless Interfaces between People, Bits and Atoms. *Proc. of CHI'97*.
- 4 Römer, A. (2002). *Unterstützung des Design Problem Solving: einsatz und Nutzen einfacher externer Hilfsmittel in den frühen Phasen des konstruktiven Entwurfprozesses*. Dresden, Dissertation, TU Dresden.
- 5 Tversky, B. (2003). *Sketching for Design and Design of Sketches. Human Behaviour in Design*. Berlin, Springer-Verlag.