

# Kinect

## als Eingabegerät



# Inhaltsverzeichnis

1	Einleitung.....	3
2	Kinect.....	4
2.1	Technische Daten .....	4
2.1.1	Sehfeld.....	4
2.1.2	Auflösung.....	4
2.1.3	Audio .....	4
2.1.4	Skeletal Tracking System .....	4
2.2	3D Kamera .....	5
2.2.1	Time of Flight .....	5
3	Kinect am Computer.....	6
3.1	Einsatzbereiche.....	6
3.1.1	Präsentationen .....	6
3.1.2	Multimediaanwendungen .....	6
3.1.3	Grafik Software.....	6
3.1.4	Motion Capture .....	6
3.1.5	In der Informatik .....	7
3.1.6	Robotik .....	7
3.2	Grenzen der Kinect .....	7
3.2.1	Erkennbare Gelenke .....	7
3.2.2	Auflösung.....	7
3.2.3	Sichtfeld.....	8
3.2.4	Licht .....	8
4	Programmier Plattformen .....	9
4.1	OpenKinect .....	9
4.2	CL NUI .....	9
4.3	Microsoft SDK .....	9
5	Fazit .....	10
6	Literatur.....	11

# 1 Einleitung

Die vorliegende Ausarbeitung entstand im Rahmen des Hauptseminars „Software für technische Systeme“ mit dem Thema „Kinect als Eingabegerät“. Sie befasst sich mit der Steuerung von Computer- und Multimediageräten mittels der Kinect. Ziel der Ausarbeitung ist es über mögliche Einsatzgebiete in denen die Benutzung einer Kinect von Vorteil wäre zu informieren. Zudem auch Probleme und schlechte Einsatzgebiete, für die die Kinect weniger geeignet ist aufzuzeigen. Zum Testen wurden verschiedene Plattformen aus freien Projekten genutzt.

# 2 Kinect

Die Kinect wurde erstmals unter dem Codenamen Projekt Natal, auf der E3-Spielemesse, vorgestellt. Das Referenzdesign für die Kinect lieferte die Firma PrimeSense. Ursprünglich war sie zur Steuerung von Spielen durch Gestenerkennung gedacht. Bis heute wurde die Kinect über 10 Millionen Mal verkauft. Am 3. Januar 2011 wurde ein offizieller Klon für den PC vorgestellt, der Asus Wavi Xtion.

## 2.1 Technische Daten

Folgend ein Überblick über die technischen Eigenschaften der Kinect. Das Gerät verfügt über eine VGA-Kamera für ein normales Bild, einen Tiefensensor zur Erfassung der Entfernung zu einem Objekt und ein Mikrofon-Array.

### 2.1.1 Sehfeld

Das Sehfeld der Kinect ist im Horizontalen  $57^\circ$  und im Vertikalen  $43^\circ$ . Zusätzlich besitzt sie einen Motor, der die Kinect um  $27^\circ$  neigen kann. Anhand dieser Zahlen ist schon zu erkennen, warum ein Mindestabstand zu der Kamera einzuhalten ist. Steht eine Person zu nah am Gerät ist der vollständige Körper nicht im Sichtfeld der Kinect. Besonders im Vertikalen, da Menschen im üblichen auch größer als breit sind. Zudem ist der Bereich in dem der Tiefensensor verlässliche Werte liefert eingeschränkt, das ist auf die grundsätzliche Funktionsweise dieser Technologie zurückzuführen. Laut Herstellerangaben soll die Kinect in einem Bereich von 1,2 bis 3,5 Metern zuverlässig arbeiten. Dies gilt aber nur solange eine komplette Person erkannt werden soll. Will man nur die Arme einer Person erfassen ist ein geringerer Abstand zum Gerät möglich.

### 2.1.2 Auflösung

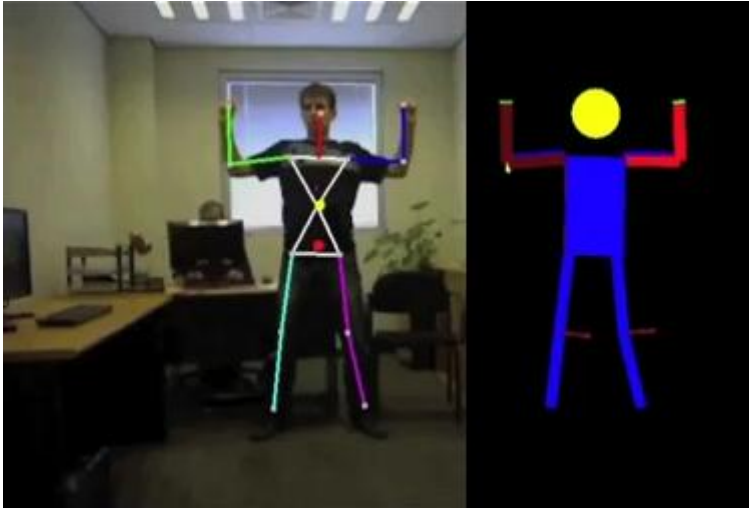
Laut Wikipedia liefern die Optischen Sensoren der Kinect Bilder mit einer Auflösung von  $640 \times 480$  Pixeln. Die Plattformen die in freien Projekten entstanden sind liefern das Bild des Tiefensensors in einer geringeren Auflösung, da es noch Probleme mit den Treibern gibt. Bei dieser geringen Auflösung ist schon zu erkennen, dass Details wie Finger für die Kinect schwer zu erkennen sind.

### 2.1.3 Audio

Ein Mikrofon-Array aus 4 Mikrofonen mit einer Abtastrate von 16kHz. Durch die vier Mikrofone kann ermittelt werden aus welcher Richtung ein Ton kommt.

### 2.1.4 Skeletal Tracking System

Die Kinect kann zwei aktive Nutzer erkennen, jedoch allgemein so viele wie in das Sichtfeld passen. Pro aktivem Nutzer kann das Gerät 20 Gelenke im "Skelett Modell" erstellen. Das Verfahren dazu nennt sich „skeletonization“. Dabei verfolgt die Kinect die Körper der aktiven Nutzer und stellt sie als Serie von Gelenken im Raum dar. So können die Positionen der einzelnen Gelenke einfach abgerufen werden.



Auf dem Bild (links) ist das VGA Bild zu sehen mit den Gelenken die erkannt wurden und (rechts) ist ein Modell zusehen, dass aus den Gelenkdaten errechnet wurde.

## 2.2 3D Kamera

Um insbesondere Probleme die bei der Verwendung mit der Kinect vorkommen zu verstehen, soll an dieser Stelle kurz erklärt werden wie eine 3D Kamera funktioniert.

### 2.2.1 Time of Flight

Das Time-of-Flight Lichtzeitverfahren (TOF) ist eine bekannte Methode zur Entfernungsbestimmung, das vereinfacht nach folgendem Prinzip arbeitet. Licht breitet sich abhängig vom Medium mit einer endlichen Geschwindigkeit aus. Deswegen benötigt es eine Gewisse Zeit um von einem Sender zu einem Empfänger zu gelangen. Somit kann unter Berücksichtigung der zweifach zurückgelegten Distanz die Entfernung bestimmt werden. Bedenkt man jedoch das Licht 1cm in einer Zeit von etwa 33ps zurücklegt, wird offensichtlich das eine genaue Zeitmessung unerlässlich für eine zuverlässige Entfernungsmessung ist. Das macht die Umsetzung jedoch sehr aufwendig und kostenintensiv.

### 2.2.2 Modulation

Um Ressourcen und damit Kosten zu sparen achtet man auf die Phasenverschiebung. Diese wird wie folgt berechnet:

$$I = a_0 + a_1 \cos(\alpha + \Delta\phi)$$

Offset  $a_0$  (Hintergrundlicht)

Amplitude  $a_1$  (Reflektiertes moduliertes Licht)

Phasenverschiebung  $\Delta\phi$

Für die drei Unbekannten werden 3 Messungen benötigt, jedoch wird oft noch eine zusätzliche Messung durchgeführt um ein robusteres Ergebnis zu erhalten.

# 3 Kinect am Computer

Wozu braucht man eine Kamera als Eingabegerät? Maus und Tastatur haben bisher für die Benutzung von Computern ausgereicht. Welche Vorteile soll die neue Steuermöglichkeit bringen und kann sie Maus und Tastatur ersetzen oder nur ergänzen? Im Folgenden werden mögliche Einsatzgebiete beschrieben. Außerdem auch Bereiche in denen der Einsatz einer Kinect wenig Sinn macht oder Probleme die der Einsatz der Kinect mitsichbringt.

## 3.1 Einsatzbereiche

Wo kann man die Vorteile der Kinect ausnutzen und wo bringt sie Vorteile im Vergleich zu Maus und Tastatur.

### 3.1.1 Präsentationen

Bei Präsentationen kann die Steuerung durch Gesten mehr Flexibilität und Bewegungsfreiheit verschaffen. Die Präsentation muss nicht mehr von stationär gelegener Maus oder Tastatur gesteuert werden. Stattdessen kann sie von nahezu überall gesteuert werden, nur eingeschränkt durch das Sichtfeld der Kinect. Probleme könnten hier ungewollte Befehle durch die normalen, ans Publikum gerichteten Gesten bereiten.

### 3.1.2 Multimediaanwendungen

Eines der klassischsten Anwendungsfälle von gestengesteuerten Programmen sind Multimediaanwendungen. Viele TV Hersteller haben bereits Geräte mit integrierter Kamera zur Gestenerkennung angekündigt oder gar schon auf den Markt gebracht. So kann man das Fernsehprogramm, Filme und Musik durch Gesten steuern. Das bringt für viele Privatanwender viel Komfort, da sie ihr Gerät vom Sofa aus steuern können. Jedoch wäre eine zentrale Kamera mit der mehrere unterschiedliche Geräte gesteuert werden können von Vorteil.

### 3.1.3 Grafik Software

Grafik Software bietet sich an um mit Gesten gesteuert zu werden, insbesondere 3D-Modellierung. Da man durch die Maussteuerung an zwei Dimensionen gebunden ist wäre die Steuerung durch eine 3D-Kamera naheliegend. Modelle wären zudem mithilfe eines 3D-Bildschirms "berührbar".

### 3.1.4 Motion Capture

Mit der Kinect ist auch ein Echtzeit-Motion-Capture möglich. So lässt sich das erstellte Skelet einer Person auf verschiedene Figuren übertragen. Das würde bedeuten, dass jeder mit geringen Kosten Motion Capture für kleine Animationen, Spiele oder ähnliches erstellen kann.

### **3.1.5 In der Informatik**

Informatiker könnte man die Kinect bei der UML Modellierung nutzen. Da das Verschieben von Objekten auf einer Fläche eine gut zu realisierende Anwendung wäre.

### **3.1.6 Robotik**

In der Robotik könnte man zum einen die Kinect dazu verwenden den Roboter fernzusteuern. Besonders sinnvoll wäre dies für Roboter die einen menschlichen Arm nachbilden.

Eine weitere Anwendungsmöglichkeit wäre einen Roboter mit einer Kinect auszustatten um ihm das „sehen“ zu ermöglichen. Es wäre auf kostengünstige Weise möglich, dass der Roboter seine Umgebung 3-Dimensional erfassen kann um somit einfacher um Hindernisse herumz navigieren zu können.

## **3.2 Grenzen der Kinect**

Natürlich ist noch nicht alles möglich, insbesondere weil die Kinect Hardwareseitig noch nicht auf die Computersteuerung ausgelegt ist.

### **3.2.1 Erkennbare Gelenke**

Bei der Benutzererkennung gibt es noch einige Grenzen die es sehr erschweren die Kinect als einziges Eingabegerät für Computer zu nutzen. Zum einen kann das Gerät momentan laut Hersteller Angaben nur 20 Gelenke erkennen, was bei einer virtuellen Tastatur zu Schwierigkeiten führen würde. Da eine Hand 15 Gelenke hat - zwar muss man nicht alle Gelenke erfassen um die Bedienung einer Virtuellen Tastatur zu realisieren - jedoch selbst wenn man nur 5 Gelenke pro Hand erfassen würde (jeder Finger wird durch ein Gelenk dargestellt), hätte man allein mit den Händen schon die Hälfte der Kapazität erreicht.

### **3.2.2 Auflösung**

Ein weiterer Punkt ist die Auflösung, die Kinect kann Finger bei dem vorgeschriebenen Abstand zum Gerät überhaupt nicht erkennen (laut Auflösungsangaben auf Wikipedia). Dadurch wäre die Flexibilität wieder eingeschränkt. Bei geringerem Abstand zur Kinect ist die Erkennung von Fingern möglich, jedoch wird dann auch nicht der ganze Körper des Nutzers erkannt.



Die Abbildung zeigt (links) das VGA Bild eines Raums und (rechts) das Bild des Tiefensensors. Man sieht sofort, dass die Auflösung des Bilds relativ gering ist und im hinteren Bereich des Raumes wenig Details zu erkennen sind.

### **3.2.3 Sichtfeld**

Das Sichtfeld der Kinect ist besonders im Horizontalen eingeschränkt, da sie über einen Motor verfügt der das Vertikale Sichtfeld erweitert. Das führt dazu, dass man zum Beispiel bei Präsentationen aus dem Sichtfeld der Kinect läuft.

Im Vertikalen führt der geringe Winkel dazu, dass man einen Mindestabstand einhalten muss sofern man den kompletten Körper einer Person erfassen möchte.

### **3.2.4 Licht**

Die Kinect ist sehr empfindlich gegen Fremdlicht. Bei zu viel Sonnenlicht liefert die Kamera zu viele Fehlmessungen. Daraus folgt, dass die Kinect nur Räumen genutzt werden kann die passend beleuchtet sind und unter freiem Himmel gar nicht.



# 4 Programmier Plattformen

Anschließend sollen noch einige Plattformen genannt werden, die die Benutzung der Kinect am Computer ermöglichen. Auch das SDK von Microsoft das am 16. Juni erschienen ist soll kurz erwähnt werden, auch wenn es nicht während dieser Arbeit getestet wurde.

## 4.1 OpenKinect

([http://openkinect.org/wiki/Main\\_Page](http://openkinect.org/wiki/Main_Page))

Die OpenKinect Community befasst sich damit die Kinect am Computer und anderen Geräten nutzbar zu machen. Durch die große Zahl der Community-Mitglieder entwickelt sich das Projekt schnell weiter und man wird ständig über Neuigkeiten informiert.

Ein Nachteil ist, dass aus der Größe des Projekts auch eine gewisse Komplexität gegeben ist. Die Treiber haben nicht korrekt funktioniert und man muss einen relativ hohen Zeitaufwand in die Einarbeitung der Plattform einrechnen. Positiv hervorzuheben ist das man eine größere Freiheit bei der Sprachwahl hat. Die Entwicklung mit der OpenKinect Plattform ist unter anderem in verschiedenen C-Sprachen, JAVA und Python möglich.

## 4.2 CL NUI

(<http://codelaboratories.com/nui>)

Die Plattform CL NUI wird von einer weitaus kleineren Gruppe entwickelt und wird nicht von einer Community entwickelt der man jederzeit beitreten kann. Jedoch sind die Ergebnisse Open Source und freizugänglich. Das mitgelieferte Testprogramm funktioniert nach dem Installieren sofort und Problemlos. Ein großer Nachteil der CL NUI Plattform ist jedoch, dass das letzte Update im Dezember 2010 veröffentlicht wurde. Die Plattform hat auch noch einige Einschränkungen in ihren Möglichkeiten, so ist zum Beispiel das Mikrofonarray nicht nutzbar.

## 4.3 Microsoft SDK

(<http://research.microsoft.com/en-us/um/redmond/projects/kinectsdk/>)

Das Microsoft SDK bietet im Vergleich zu den anderen Plattformen vollen Zugriff auf alle Funktionen der Kinect. So kann man sowohl auf die unveränderten Daten der beiden Kameras, also dem Tiefensensor und der VGA-Kamera, als auch von sämtlichen Mikrofonen, zugreifen.

Ein Nachteil des Microsoft SDKs ist das man sich auf .NET Sprachen einschränken muss.

# 5 Fazit

Zwar hat die Gestensteuerung noch mit einigen Schwierigkeiten zu kämpfen, jedoch gibt in der Theorie genügend sinnvolle Anwendungsgebiete, dass es sich lohnt an ausgereiften Softwareprodukten für die Kinect zu arbeiten. Besonders da die Hardware, also die Kinect-Kamera selbst schon sehr weit verbreitet ist.

Dabei dürfen die Probleme aber nicht außer Acht gelassen werden. In Räumen mit viel störendem Fremdlicht oder gar unter freiem Himmel wird die Kinect wohl nie einsetzbar sein.

Jedoch sind Probleme wie Auflösung und die Anzahl der Gelenke möglicherweise lösbar oder zumindest eindämmbar. Besonders die Anzahl der Gelenke sollte Softwareseitig gut erweiterbar sein. Somit wären auch Lösungen möglich, durch die man die Tastatur, mit der Kinect ersetzen oder zumindest bei kurzen Texten ablösen könnte. Allerdings wäre weiter zu beachten, dass man einen geringeren Abstand zum Gerät einhalten muss da die Auflösung weiterhin nicht ausreichen ist um einzelne Finger auf große Distanz zu erkennen. Das Problem der geringen Auflösung der Kinect lässt sich zudem auch nur bedingt Softwareseitig beheben. Bestenfalls kann man die Bilder noch etwas digital nachbearbeiten. Da das Windows SDK erschienen ist kann man hoffen, dass dadurch bessere Bilder geliefert werden, da die freien Plattformen Probleme haben ein optimales Bild zu erhalten.

Abschließend ist zu sagen, dass die Kinect noch nicht sofort einsetzbar ist. Jedoch wurde den Entwicklern mit dem Microsoft SDK neue Möglichkeiten eröffnet. Also bleibt abzuwarten ob die Kinect eine neue Ära der Bedienkonzepte einleitet. Vielleicht bringt auch die Lösung von Asus ausschlaggebende Neuerungen und sticht den Konkurrenten aus.

## 6 Literatur

1. Dipl.-Ing- Holger Kraft "Untersuchung und Entwicklung integrierbarer Photomischdetektor (PMD)-Konzepte auf Halbleiterbasis zur Realisierung hochauflösender 3D-Messsysteme"
2. Bernd Jähne "Entwicklungsstand und Zukunftsperspektiven für 3D-Laufzeitkameras"
3. <http://en.wikipedia.org/wiki/Kinect>
4. <http://gizmodo.com/5682023/microsoft-kinect-gutted-four-microphones-three-cameras-and-one-very-important-fan>
5. <http://www.hardware-infos.com/news.php?news=3766>