

# 1 Kinect – was ist das?

Wenn Sie dieses Buch in Händen halten, wissen Sie sehr wahrscheinlich, dass es sich beim Kinect um ein von Microsoft vertriebenes Eingabegerät für Computersysteme handelt.

Aber: Wissen Sie auch, wie die einzelnen Hardwarekomponenten funktionieren? Wissen Sie, was das Projekt Natal war? Oder haben Sie sich schon immer gefragt, ob der Kinect in der Wiener U-Bahn beworben wurde?

Betrachten Sie dieses erste Kapitel nicht als Lehrstoff, sondern als Einführung in die Thematik. Wenn Sie diese Einführung aus Zeitmangel überspringen wollen, lesen wir uns später wieder. Wer wie ich Nebensächlichkeiten liebt, ist hier genau richtig.

## 1.1 Entfernungen messen

Klassische Entfernungssensoren für Gestenerkennung arbeiteten mit zwei bekannten Schemata: Kontrasterkennung und Signallaufzeit (Time-of-flight).

Die Schwächen der Kontrasterkennung lassen sich an Abbildung 1–1 ablesen. Die graue Farbe des Hemds ist kaum von der weißen Leinwand im Hintergrund zu unterscheiden. Ohne Tiefeninformationen (sprich nur mit Kontrasterkennung) ist es praktisch unmöglich, den Torso des Autors vom Hintergrund zu lösen. Zudem lässt sich der Abstand zum Sensor anhand der Farben allein nicht bestimmen. Auch wenn die Wand 20 Meter von der Kamera entfernt ist, bleibt sie weiß und wird nicht blau – obwohl diese Vorstellung in der Antike weit verbreitet war.

Anders als die Kontrasterkennung liefert die Time-of-flight-Methode echte Tiefendaten. Dabei kommt eine auf den ersten Blick primitive Methode zum Einsatz: Eine Lichtquelle feuert Infrarot- oder Lasersignale ab, die vom Objekt reflektiert und von einem Sensor eingefangen werden.



**Abb. 1–1**      *Weiß auf grau: das Ende der Kontrasterkennung*

Anhand der Signallaufzeit kann man sodann auf die Entfernung des reflektierenden Punkts schließen. Diese auf den ersten Blick primitive Technologie ist nicht ohne, da der durchschnittliche CMOS-Sensor nicht in der Lage ist, Laufzeiten zu ermitteln.

Der geringe Preis der Hardware wurde erst durch eine Entwicklung von PrimeSense möglich. Der Kinect ist – man mag es angesichts der enormen Marketinganstrengungen von Microsoft gar nicht glauben – streng genommen keine Entwicklung aus Redmond. Die im Sensor zum Einsatz kommende Technologie ist geheim. Es gibt außerhalb der diversen Patente kaum offizielle Beschreibungen, wie das System tatsächlich arbeitet.

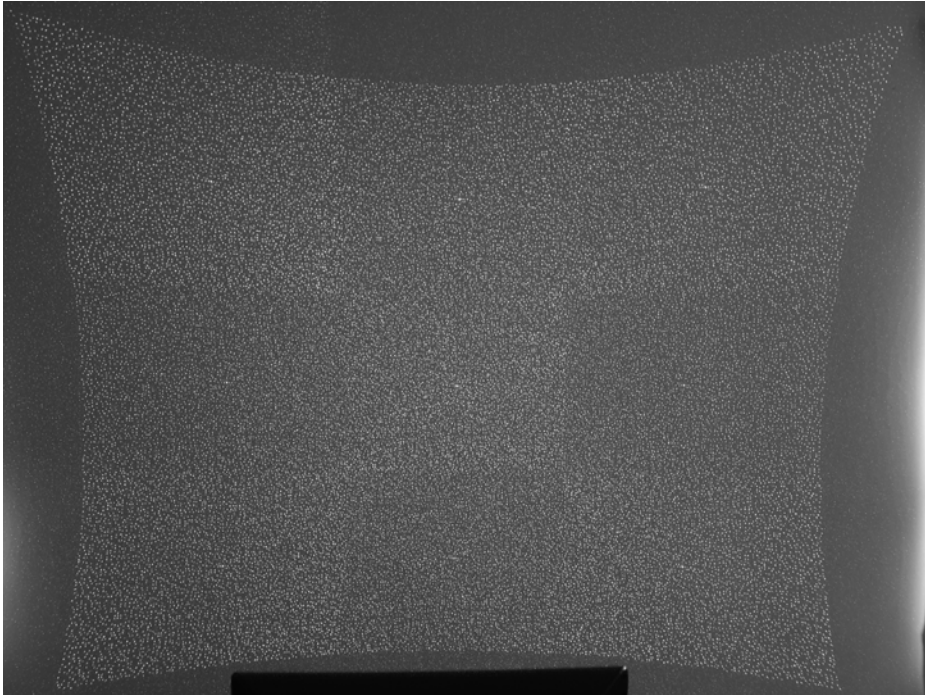
Anhand der verfügbaren Informationen hat John MacCormick vom amerikanischen Dickinson College eine sehr interessante Präsentation erstellt, die kostenlos (da mit amerikanischem Steuergeld erstellt) heruntergeladen werden kann.<sup>1</sup>

Die Grundidee ist einfach: Der Kinect sendet das in Abbildung 1–2 gezeigte Punktemuster in die Umwelt. Das Punktemuster durchquert eine spezielle Linse, die die ausgehenden Lichtstrahlen deformiert. Diese Deformation hat den Effekt, dass sich die Form der einzelnen Punkte je nach dem Abstand zum Sensor verändert.

Der Kinect muss die Infrarotszene fotografieren und die Infrarotkügelchen erkennen. Anhand ihrer Form kann er danach auf die Entfernung der einzelnen Objekte schließen. Das Resultat davon ist eine Tiefenmatrix.

---

1. <http://users.dickinson.edu/~jmac/selected-talks/kinect.pdf>



**Abb. 1-2** Die Kinect-Punktwolke (Quelle: <http://www.futurepicture.org/?p=129>)

Microsoft bestätigt in MSDN (Microsoft Developer Network), dass der Kinect für die Tiefenerkennung Infrarot verwendet. Es ist also sehr wahrscheinlich, dass die hier gegebene Beschreibung zumindest annähernd korrekt ist.

## 1.2 Doing it the Microsoft Way

Die Entwicklung der Time-of-flight-Methode war für sich genommen noch kein besonderes Ereignis. Wirklich interessant wurde die Technologie erst, als die vier Gründer von PrimeSense mit einem Prototyp ihres Geräts zu Microsoft pilgerten.

Dort fand man sehr schnell Interesse an der Technologie. Am 30.5.2007 sprach Bill Gates auf der D5-Konferenz beiläufig von einer »Kamera zur Spielesteuerung«.

Aus damaliger Sicht war dies keine besondere Ankündigung. Kameras zur Spielesteuerung gab es schon damals für verschiedene Konsolen. Als Beispiel sei hier die 2003 für die PlayStation ausgelieferte EyeToy-Kamera genannt, oder auch die heute kaum noch bekannte U-Force-Kamera für das Nintendo Entertainment System, die sogar schon mit Infrarot arbeitete.<sup>2</sup>

---

2. <http://en.wikipedia.org/wiki/U-Force>

Der Erfolg der Nintendo Wii führte am 7.4.2008 zu einem weiteren »Aufflackern« in der Gerüchteküche: MTV News verlautbarte, dass Microsoft die Nintendo'sche Konsole angreifen wolle. Wie das erfolgen sollte, blieb damals offen.

Erst die Spielemesse E3 brachte 2009 weitere Einblicke. Microsoft kündigte auf seiner alljährlichen Pressekonferenz das Project Natal an. Das Project Natal wurde als »bewegungsempfindlicher« Sensor für die Xbox vorgestellt. Vom Namen Kinect wusste man damals noch nichts.

Für Entwickler war die Verlautbarung im Januar 2010 mit Abstand am wichtigsten. Microsoft kündigte an, dass der Kinect – anders als die diversen Prototypen und Systeme von PrimeSense – keinen integrierten Prozessor haben würde. Die Berechnungen sollten stattdessen von einem der Kerne der Xbox 360 erledigt werden, deren Gesamtleistung durch diese Zusatzbelastung um rund 10% »sinken« würde.

### 1.3 Reverse Engineering zugunsten von Unix

Die Auslieferung des Kinect am 4. November 2010 war ein signifikanter Erfolg für Microsoft. Das Gerät verkaufte sich wie geschnitten Brot. Es tauchten sogar in der Wiener U-Bahn Werbestände für das Produkt auf.<sup>3</sup>

Leider veröffentlichte Microsoft bei der Erstauslieferung der Hardware keinerlei Pläne für PC-Treiber. Dadurch sah sich das für Elektronikbasteleien bekannte Unternehmen Adafruit Industries inspiriert, eine Prämie von 1000 US-Dollar für den ersten funktionierenden Treiber auszuloben.

Microsoft reagierte darauf mit einer Ankündigung, gegen jede Form von Hacking gerichtlich vorzugehen. Adafruit verdoppelte daraufhin die Belohnung, die Technikpresse schrie auf – und die Sache entpuppte sich als großes Missverständnis.

Microsoft ging es primär um Raubkopien und »Cheaten« in Multiplayerspielen. Solange es nur um den Zugriff auf die vom Sensor gelieferten Daten ging, war das dem Unternehmen egal. Das USB-Interface wurde vielmehr als »offen« beschrieben.

Auch in Redmond war man nicht faul. Das im Juni 2011 erstmals ausgelieferte offizielle SDK wurde sogar von einem eigenen, für den PC-Einsatz zugelassenen Kinect begleitet, den Sie im nächsten Kapitel kennenlernen werden.

An dieser Stelle ist nur eine Sache wichtig: Es gibt zwei konkurrierende SDKs. Das eine stammt von Microsoft, das andere ist quelloffen. Für das eine System entwickelte Anwendungen sind mit dem jeweils anderen inkompatibel.

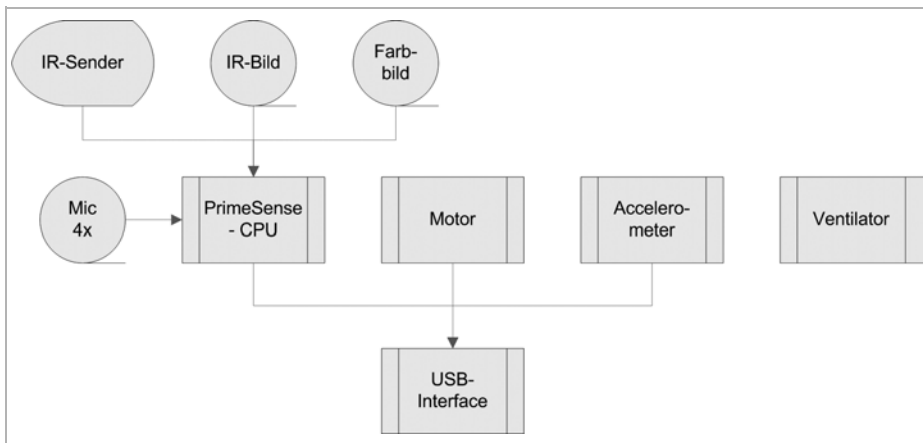
---

3. <http://tamspcc.tamoggemon.com/2011/12/01/microsoft-kinects-at-austrian-underground-rail/>

## 1.4 Innereien

Die innere Struktur des Kinect ist bei allen Modellen gleich. Eine vom amerikanischen Unternehmen iFixit veröffentlichte schrittweise Zerlegungsbeschreibung erlaubt uns, diese näher anzusehen.<sup>4</sup>

Anstatt die Hauptplatinen einzeln abzdrukken (ein Blick auf die URL <http://www.ifixit.com/Teardown/Microsoft+Kinect+Teardown/4066/1> lohnt sich für Interessierte), soll hier das in Abbildung 1–3 gezeigte Flussdiagramm ausreichen. Es zeigt schematisch alle Baugruppen des Kinect.



**Abb. 1–3** Der Kinect als Flussdiagramm; die Stromversorgung ist nicht abgebildet.

Die Kommunikation zwischen Sensor und Host erfolgt über USB. Dass der Steckplatz nicht unbedingt standardisiert aussieht, wird im nächsten Kapitel besprochen.

Neben dem zur Kühlung eingebauten (und vom PC aus nicht steuerbaren) Ventilator gibt es an »simpler« Peripherie zusätzlich ein Accelerometer und einen Motor zur Anpassung der Sichtachse.

Die eigentliche Intelligenz befindet sich in der abgespeckten PrimeSense-CPU. Sie hat die Aufgabe, die verschiedenen im Kinect enthaltenen Sensoren miteinander abzustimmen und als Resultat einen Audio-, einen Tiefen- und einen Farbdatenstrom an den Host zurückzuliefern.

Die Aussage von Microsoft, der Kinect hätte KEINEN Prozessor, ist so zu verstehen, dass die Skelettverfolgung (also der Prozess der Umwandlung von Farb- und Tiefendaten in Skelettkoordinaten) ausschließlich durch den Host erfolgt.

4. <http://www.ifixit.com/Teardown/Microsoft-Kinect-Teardown/4066/>

## 1.5 Fazit

Die vom Kinect gebotenen Interaktionsmöglichkeiten sind – per se – nichts Neues. Die Stärke des Produkts liegt im geringen Preis und der hohen Verfügbarkeit. Nie zuvor gab es ein so leistungsfähiges und gleichwohl günstiges Sensorpaket in einem so praktischen Formfaktor.