

DKE-Projekt Painect – Gesten und Posen

Unterschied Geste – Pose

Bei der Kinect oder einer ähnlichen Lösung als Eingabegerät unterscheidet man zwischen Gesten und Posen. Dabei entsprechen Gesten einem kompletten Bewegungsablauf und Posen lediglich der Positionierung des Körpers oder Körperteilen, um eine bestimmte Aktion auszuführen.

Posen lassen sich wesentlich einfacher implementieren, da hier nur boolsche Abfragen der aktuellen Positionen benötigt werden. Bei Gesten hingegen müssen die Positionen in zeitlichen Zusammenhang (z.B. über eine Sekunde) betrachtet und bewertet werden.

Eine Möglichkeit hierfür wäre zum Beispiel die Speicherung der Koordinaten in einem Array und der Betrachtung des Verlaufs der jeweiligen Punkte. Dies hat jedoch eine erhebliche Verzögerung zur Folge und bietet sich daher hauptsächlich für Eingaben an, bei denen der zeitliche Aspekt (Verzögerung etc.) eine eher untergeordnete Rolle spielt (Ausnahme: Die von Microsoft selbst für Kinect-Spiele eingesetzten Gesten, die allerdings mit Millionen von Probandendaten in großen Server-Clustern optimiert wurden...).

Ein weiteres Problem besteht darin, dass eine gewisse Toleranz bei der Gestenerkennung beachtet werden muss, da die Bewegungen der Nutzer untereinander und im Vergleich zur implementierten Geste nie identisch (bzgl. zeitlichem Ablauf und Positionen) sein können.

Keep it simple

Grundlegend kann gesagt werden, dass man es dem Anwender ebenso wie sich selbst als Entwickler einen Gefallen tut, wenn man auch die Gesten oder Posen einfach hält.

Komplexe Bewegungen verlangen dem Nutzer ein wesentlich höheres Maß an Koordination und Genauigkeit bei der Ausführung ab. Zudem mindert eine zu komplexe Steuerung die intuitive Bedienung der Applikation und kann für Frustration beim Anwender sorgen, wenn Steuerungsversuche häufig scheitern.

Als Programmierer erfordern komplizierten Gesten und Posen einen sehr viel größeren Implementations- und Testaufwand, damit das Endergebnis auch wirklich der Vorstellung entspricht.

Swipe-Gesten und Lean-Posen

Swipe-Gesten, die durch boolsche Abfragen (z.B., ob der Abstand zwischen aktuellem Punkt und Referenzpunkt einen bestimmten Wert überschreitet) bedürfen für den Anwender einiges an Übung, da der Referenzpunkt sehr statisch ist. Zu hektische Bewegungen führen dazu, dass es zu Problemen bei der Erkennung kommen kann und dabei Gesten zu häufig oder gar nicht erkannt werden.

Ein ähnliches Problem besteht bei Posen, bei denen man den Oberkörper neigt oder dreht. Für diese muss ebenfalls eine Referenz-/Ausgangsposition definiert werden. Zwar ist es möglich, die Ausgangsposition in bestimmten Intervallen zu aktualisieren, allerdings ist damit ein dauerhaftes Ausführen der Posen-Aktion nicht mehr möglich. Die Pose würde nach Ablauf der Intervallzeit selbst als Ausgangsposition definiert werden und somit Fehler in der Posen-Erkennung auftreten.

Diese Probleme lassen sich vermeiden, indem auch diese Gesten und Posen anhand des Bewegungsablaufs getriggert werden. So müsste ein zur Seite Neigen z.B. anhand des Bewegungsablaufs bestimmter Punkte und nicht durch deren aktuelle Position im Vergleich zu ihrer Ausgangsposition erkannt werden. Daraus folgen jedoch die im ersten Abschnitt („Unterschied Geste – Pose“) beschriebenen Probleme.

Überlagerung von Gesten

Bei der Auswahl von Gesten und Posen für ein Kinect-Projekt ist ebenso die Überlagerung von Gesten zu vermeiden. Zum einen dürfen sich Gesten und Posen nicht bei ihrer direkten Ausführung überlagern (z.B. ist eine klare Trennung zwischen Hand heben und gesamten Arm heben zu sichern, wenn beide Posen/Gesten genutzt werden sollen), zum anderen darf beim Einnehmen einer Pose keine andere Pose oder Geste ausgelöst werden. So kann das gleichzeitig Verwenden von Posen für das seitliche Ausstrecken eines Armes in 90° und 45° zum Körper zu diesem Problem führen.

Somit sind innerhalb der Applikation oder zumindest innerhalb der Modi eindeutig unterscheidbare Gesten und Posen zu verwenden oder ggf. klarere Differenzierungen bei der Abfrage einzuführen.

Feintuning von Gesten

Das Feintuning von Gesten ist ohne entsprechend große Anzahl an Testdaten (siehe Beispiel Microsoft aus dem Abschnitt „Unterschied Geste – Pose“) durchaus kompliziert, weshalb hier nur einige Tipps und Hinweise gegeben werden können.

Zu beachten sind sowohl Abstand und Position zur Kamera, als auch Größe und Proportionen der Probanden. Dabei sollte man sowohl den Abstand zur Kamera variieren, als auch mit verschiedenen Probanden (Kinder, Frauen, Männer, große und kleine Personen etc.) testen, um eine angemessenes Ergebnis zu erhalten.

Einen enormen Vorteil erhält man durch die Nutzung von dynamischen Längenmaßen. Nimmt man z.B. den Abstand zwischen Hand und Ellenbogen als Berechnungsgrundlage für Abstände, so kann man dadurch die Abhängigkeiten von Gesten bzgl. Abstand und Position zur Kamera minimieren und zudem teilweise auf die unterschiedlichen Körpermaße der Nutzer reagieren.

Für Gesten, die in einem zeitlichen Verlauf ausgeführt werden, spielt der Zeitraum für die Erkennung eine große Rolle. Einerseits darf der Zeitraum nicht zu klein sein, damit die Geste problemlos ausgeführt werden kann, andererseits sollte der Zeitraum auch nicht zu groß gewählt werden, da die Handhabung der Applikation sonst erheblich durch die Verzögerung eingeschränkt wird.

Wie im zweiten Abschnitt („Keep it simple“) bereits angesprochen, ist es auch beim Feintuning hilfreich, wenn die Gesten und Posen einfach gehalten wurden. Die Fehlerquelle Anwender wird somit auf ein Minimum reduziert. Eine unsaubere Ausführung der Gesten erschwert die Erkennung bereits bei einfachen Gesten. Je komplexer die Gesten werden, desto größer ist auch die Gefahr, dass die Ausführung zu ungenau wird.

Empfehlung

Während der Entwicklung von Painect haben sich einige Posen und Gesten als geeigneter herausgestellt und andere zu größeren Problemen geführt. Allerdings nutzt das Projekt (bisher) eher statische Abfragen und damit eher Posen als Gesten im eigentlichen Sinne. Die folgende Liste ist eine subjektive Bewertung der Posen und Gesten entsprechend Ausführbarkeit und Implementationsaufwand.

Geste	Ausführbarkeit	Implementationsaufwand
Gut		
Arm zur Seite ausstrecken	Einfach	Boolsche Abfrage → Einfach
Arm nach vorn ausstrecken	Einfach	Boolsche Abfrage → Einfach
Arm nach oben strecken	Einfach	Boolsche Abfrage → Einfach
Hand heben	Einfach	Boolsche Abfrage → Einfach
Alle vorher genannten Gesten in beidseitiger Ausführung	Einfach	Abfrage beider Seiten gleichzeitig → Einfach
Mittelmäßig		
Swipe-Gesten senkrecht, waagerecht, diagonal auf einer Ebene (z.B. x-y-Ebene)	Schwierigkeiten, wenn Referenzpunkt beachtet werden muss	Mittels Vergleich zwischen aktueller Position und Referenz → Einfach, aber mehr Probleme bei der Ausführbarkeit Mittels Analyse in best. Zeitfenster → Komplizierter
Schlecht		
Lehnen/Neigen des Oberkörpers	Extreme Winkel komplizierter	Hoher Aufwand, da verlassen der Ausgangsposition beachtet werden muss
Seitliches Drehen		
Laufen und Springen	Je nach Komplexität	Kompliziert, da sowohl die Erkennung als auch die Abgrenzung von anderen Gesten gewährleistet werden muss
Allg. komplizierte Bewegungsabläufe		