Laboratorium CAD/CAM

Rozpoznawanie obiektów 3D przez sensor Kinect

Przygotowali

Michał Cwienczek

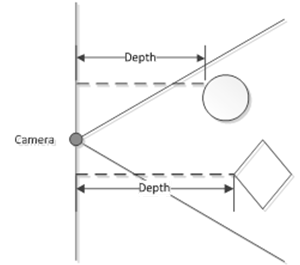
Michał Szewczak

1. Opis zadania

Zadanie polegało na przetworzeniu obrazu z kamery/kamer sensora kinect w taki sposób, aby rozpoznać obiekty znajdujące się przed sensorem. Do tego celu użyliśmy obrazu z kamer głębokości (chmura punktów głębokości) oraz obrazu z kamery zwykłej.

1. Wykonanie

Pierwszym etapem przetwarzania było pobranie chmury punków odległości z sensora kinect. Dane te są przekazywane w postaci tablicy jednowymiarowej opisujące głębokości punktów na płaszczyźnie 640x480 punktów. Każdy z tych punktów opisywany jest przez odległość w milimetrach. Sensor kinect dostarcza odległości pomiędzy płaszczyzną styczną do sensora a punktem znajdującym się przed sensorem w odległości od 0.8m do 4m (rysunek poniżej).



Następnym etapem było przetworzenie pobranej chmury punktów oraz wykrycie konkretnych obiektów na powierzchni badanej. Aby wykryć poszczególne obiekty znajdujące się przed sensorem wykorzystaliśmy różnicę w wysokościach poszczególnych pikseli. Dla danych odległości punktów x1 oraz x2 policzyliśmy

Gdzie x1 oraz x2 to odpowiednio odległość punktu x1 oraz x2 od sensora w milimetrach.

Generowaliśmy następnie spójne obszary, których delta nie przekraczała pewnego wyznaczonego progu. Okazało się iż najlepsze rezultaty osiągają małe delty, od 10 do 20 mm.

Generowanie tych wspólnych obszarów odbywało się poprzez przechodzenie chmury punktów algorytmem podobnym do BFS, lecz zmodyfikowanym na potrzeby Kinect’a.

Podczas pracy nad projektem odkryliśmy także, iż Kinect posiada obszary, dla których nie potrafi obliczyć odległości. Punkty te przyjmują odległośc równą zero. Punkty te powodowały spore utrudnienia w wykrywaniu obiektów. Aby poradzić sobie z tym, pomijaliśmy te punkty, jako, że część z nich reprezentowała wadliwe piksele. Niestety liczenie średniej z sąsiednich punktów nie dało zadowalających rezultatów, obiekty były źle kolorowane.

Pseudokod algorytmu znajduje się poniżej:

*Q – kolejka punktów*

*Dodaj do Q pierwszy punkt*

*Póki Q nie jest pusta*

*p = Q.pop();*

*dla każdego punktu pp stycznego z p*

*jeśli pp nie był odwiedzony*

*jeśli |pp – p| < delta*

*koloruj pp*

*dodaj pp to Q*

*oznacz pp jako odwiedzony*

*pokoloruj pierwszy punkt*

1. Wnioski
2. Urządzenie kinect może być wykorzystane do prostego wykrywania obiektów na pewnym bliskim obszarze w przestrzeni zamkniętej.
3. Pewna niedokładność urządzenia musi być zrekompensowana (wiele punktów miało wartość zero)
4. Dobre i wydajne przetwarzanie tablicy jednowymiarowej punktów jest kluczowa do osiągnięcia optymalnej wydajności.