

Distribuerat Mobilt Datorseende och Applikationer på Androidplattformen

Sebastian Olsson

Philip Åkesson

15 juni 2009

Bakgrund

Mobiltelefonutvecklingen har lett till att det idag är möjligt att utföra beräkningar som tidigare bara var möjliga på snabba datorer. I det här arbetet undersöker och beskriver vi implementationer av datorseendeapplikationer på Androidplattformen. Datorseende är ett samlingsnamn för metoder som tolkar och drar slutsatser utifrån bilder. Vi har koncentrerat oss på objektigenkänning; att bestämma vilka specifika objekt som förekommer i ett fotografi taget med mobiltelefonen.

Android är en öppen mobilplattform som lanserades av Google 2007 och idag utvecklas av ett antal företag i Open Handset Alliance. Den första Androidtelefonen (HTC Dream) släpptes på den amerikanska marknaden i oktober 2008, och den första på den svenska marknaden (HTC Magic) i juni 2009. Många tillverkare planerar att släppa telefoner inom en snar framtid.

Metoderna vi har använt bygger på så kallade local features. Detta bygger på att ett antal intressanta punkter hittas i bilden och omgivningen runt dessa beskrivs med tal, t.ex. histogram över intensiteten i pixlarna runt den intressanta punkten. Dessa punkter ska gå att pålitligt hitta i flera bilder tagna ur olika vinklar, ljus och avstånd. Beskrivna punkter i en bild kan sedan jämföras med en databas över kända bilder för att hitta punkter som liknar varandra.

När man utvecklar tunga och beräkningskrävande applikationer för till exempel datorseende så är det viktigt att ta hänsyn till begränsningarna hos den mobila plattformen, till exempel den låga processorhastigheten och mängden arbetsminne. Man bör samtidigt utnyttja de möjligheter som finns, som till exempel tillgång till kamera och GPS. I rapporten analyserar vi hur man bäst utvecklar den här typen av applikationer.

Resultat

Vi presenterar som exempel två applikationer skrivna i Java för att känna igen ett eller flera objekt i bilder från mobiltelefonens kamera. Vi presenterar även en tredje applikation skriven i C som använder en hårt optimerad implementation för att försöka uppnå följning av objekt i nära realtid.

Den första applikationen, Art Recognition, används för att känna igen tavlor och presentera användaren med information om tavlan och konstnären. Här görs en del av beräkningarna i telefonen, och själva matchningen mot databasen på en server.

I den andra applikationen, Bartendroid, tar användaren en bild av sitt barskåp, varefter flaskorna identifieras och presenteras tillsammans med förslag på möjliga drinkrecept. Hela bilden komprimeras och skickas till en server som utför alla beräkningar.

I den tredje applikationen, AndTracks, görs alla beräkningar i mobiltelefonen helt utan någon server. Användaren väljer först ett objekt med pekskärmen varefter igenkänningsalgoritmen körs löpande på varje bildruta i videostreamen från kameran och objektets position markeras på skärmen.

