

Populärvetenskaplig sammanfattning

Datorseende är idag ett väldigt aktivt forskningsområde och dyker upp i populära tillämpningar så som Virtual/Augmented Reality (VR/AR), självkörande bilar och autonoma drönare.

Gemensamt för dessa är att man vill utvinna tredimensionell information från tvådimensionella bilder. Detta kan t.ex. vara kamerans position och orientering, eller avståndet till något objekt som förekommer i bilderna. Detta kräver matematiska metoder och modeller. För att kunna vara användbara i verkligheten så måste metoderna vara både effektiva och robusta mot fel i indata. Exempelvis så är toleransen för fel väldigt låg då en självkörande bil ska identifiera avstånd till medtrafikanter. För VR/AR krävs att användarens position bestäms i realtid med så lite fördröjning som möjligt, då även små fördröjningar kan leda till illamående hos användaren.

I den här avhandlingen utvecklar vi metoder för att snabbt och robust kunna lösa olika sorters geometriska estimeringsproblem, som t.ex. att bestämma kamerans position relativt en 3D-modell. Denna typ av problem kan ofta beskrivas med system av polynomekvationer. Inom avhandlingsarbetet tar vi fram metoder för att effektivt kunna lösa sådana system. Metoderna vi utvecklar är generella och kan användas på många olika problem inom datorseende.

Idag bygger de flesta metoder för 3D-rekonstruktion på antagandet att världen är statisk, d.v.s. att 3D-strukturen inte förändras mellan bilderna. Men i verkligheten så förekommer också dynamiska scener, där föremålen i scenen deformerar och rör sig både relativt kameran och varandra. Till skillnad från fallet då scenen är statisk, så blir problemet väldigt svårt att modellera matematiskt då scenen är dynamisk. Om man endast kräver att rekonstruktionen ska stämma överens med bilderna så kommer det att finnas oändligt många lösningar. För att begränsa antalet lösningar så antar man att den sökta lösningen kommer från verkligheten och är enkel i något avseende. Ett vanligt antagande är till exempel att 3D punkterna ska röra sig i en jämn bana och inte hoppa omkring mellan bilderna. I den här avhandlingen arbetar vi med antagandet att 3D punkterna ska ha låg rang. Rang är ett matematiskt begrepp som mäter hur mycket beroende det finns i strukturen. En lösning med låg rang kommer att ha många punkter som rör sig på ett liknande sätt. Inom avhandlingen undersöker vi dels olika sorters rang-modeller och dels matematiska verktyg för att kunna hantera problemformuleringar där det ingår rang-villkor.