

**Utvärdering av effektiviteten hos COLMAP,
DROID-SLAM och NeRF-SLAM i
3D-rekonstruktion av vägscener**

Marcus Ascard & Farjam Movahedi

3-dimensionell (3D) rekonstruktion, ett område inom datorseende, har under senare år gjort stora framsteg som en konsekvens av de omfattande genombrött som har skett inom Artificiell Intelligens (AI). Inom datorseende har man som mål att försöka förklara hur världen ser ut för en dator, där saker som vi människor finner uppenbara behöver beskrivas i detalj med matematiska begrepp för att datorn ska få en korrekt uppfattning om omgivningen. Idag har tekniken utvecklats så långt att man kan programmera datorprogram att återskapa ett objekt eller ett område i 3D, där denna detaljerade 3D-rekonstruktion sedan kan roteras, skalas och studeras från alla håll. För att illustrera det med ett exempel så kan man tänka sig att vi skulle vilja skapa en digital modell av ett kontor. Med 3D-rekonstruktionstekniker kan vi genom att förse dessa datorprogram med en videosekvens av kontoret, samt ge programmet tillgång till inre kameraparametrar (en typ av information som är specifik för kameran som använts), därefter återskapa kontoret digitalt som en 3D modell. Utöver att återskapa miljön i 3D så gör dessa typer av program även en uppskattning om hur kameran har rört sig under inspelningen av videosekvensen.

För att öka säkerheten och förarupplevelsen i sina bilar utvecklar Volvo Cars förarassistanssystem (ADAS), så som kollisionsundvikande system, där kameror och andra sensorer som är placerade runt bilen utnyttjas för att öka bilens uppfattning av dess omgivning. Historiskt sett har man fokuserat på att använda sensorer så som ljusradarer (LiDARs) för att mäta avstånd till olika föremål i bilens omgivning, vilket likt 3D-rekonstruktion modellerar omgivningen i 3D. Detta tillvägagångssätt har däremot en del brister när målet är att återskapa en så detaljerad 3D miljö av omgivningen som möjligt. Bland annat så tillhandahåller inte LiDARs några färger, vilket försvårar urskiljning av olika objekt i 3D miljön avsevärt. Därtill är kvalitén av 3D modellen högt beroende av positionen av LiDARn. På grund av detta ville man undersöka om kamerorna på Volvobilar kan utnyttjas för att återskapa noggranna och detaljerade vägsccener i 3D, via 3D-rekonstruktionsteknikerna COLMAP, DROID-SLAM och NeRF-SLAM. Under det här examensarbetet har dessa tre tekniker använts för att återskapa 3D rekonstruktioner, med videosekvenser från simulerade miljöer såväl som verkliga bilder tagna med kamerorna från en Volvobil. För att utvärdera effektiviteten hos diverse rekonstruktionstekniker programmerades ett utvärderingsprogram.

Resultaten visade att COLMAP och DROID-SLAM kunde uppskatta kamerornas (som är monterade på Volvobilen) rörelsebanor med en felmarginal på decimeternivå. Vidare så lyckades teknikerna skapa 3D-rekonstruktioner av vägsccener med potential för att kunna användas i ADAS i framtiden. Generellt sett nådde COLMAP högst prestanda, men de långa körningstiderna för denna teknik gör den opraktisk att använda i stor skala. Den teknik som visade sig vara minst lovande för ändamålet var NeRF-SLAM, som inte lyckades producera acceptabla 3D rekonstruktioner när samtliga kameror på Volvobilen användes. Slutligen visade DROID-SLAM störst potential för detta användningsområde. Trots att algoritmen användes utan större modifikationer, visade den förmågan att generera imponerande resultat med låga körningstider. Dock krävs ytterligare forskning och finjustering för att optimera DROID-SLAMs prestanda givet Volvo Cars system.