

# Konferenssamtal med mobiltelefonen

---

POPULÄRVETENSKAPLIG SAMMANFATTNING **Elin Freberg, Emelie Skoog**

---

Med framväxten av digitala verktyg har också vårt sätt att kommunicera förändrats. Förväntningarna på kommunikationskanalerna blir större, och man hittar idag avancerade konferenssystem på de allra flesta arbetsplatser. Många av dessa system har även en funktion som kan fokusera bilden på den som talar. Detta ska förbättra användarupplevelsen för tittaren, men kräver också utrustning i form av externa kameror och mikrofoner. Detta arbete utforskar möjligheter att kunna efterlikna denna funktion i en mobilapplikation.

Att införa en mobilapplikation som kan följa talaren, gör att man som användare kan sätta upp ett mer professionellt konferenssamtal vart som helst i världen. Det som skiljer konferenssystem och mobiltelefoner åt är hårdvaran. Konferenssystem använder sig av en mikrofonarray som kan lokalisera vart ljud kommer ifrån, och på så sätt kan personen som pratar identifieras. En mobiltelefon är inte utrustad med en mikrofon som kan lokalisera ljud, vilket gör att identifieringen av vem som talar måste ske med hjälp av annan data. I litteraturen hittas bland annat ett förslag som grundar sig på tesen att den som pratar ofta gestikulerar. Det betyder att den person med mest rörelse kan identifieras som talaren. Detta innebär att det endast skulle behövas input från en kamera, vilket de flesta smartphones är utrustade med. Vidare är även smartphones utrustade med inbyggd ansiktigenkänning, som kan användas för att hitta hur många deltagare som befinner sig i bild, samt deras positioner.

Applikationen som skapades i detta arbete utgår därför ifrån de ansikten som identifieras. Från detta bestäms ett område som bestäms innefatta överkroppen. Rörelsen inom detta område mäts genom att beräkna det optiska flödet mellan varje bild. Personen vars område hade mest rörelse bestämdes vara den som talade. För att testa hur väl denna metoden fungerade jämfördes applikationens resultat med transkriberad data. Testerna visade att metoden inte fungerar tillräckligt bra för att identifiera talaren, men olika justeringar i parametrar kunde förbättra resultaten något. Även om applikationen inte var tillräckligt tillförlitlig, fanns indikationer på att en större parametersökning hade kunnat förbättra applikationen.