

EXAMENSARBETE Adaptive content-based sound compression**STUDENTER** Marcus Tedenvall, Markus Gerard**HANDLEDARE** Martin Stridh (LTH), Mersad Jelacic (Axis Communications)**EXAMINATOR** Leif Sörnmo (LTH)

Smart komprimeringsteknik för ljudinspelningar

POPULÄRVETENSKAPLIG SAMMANFATTNING **Marcus Tedenvall, Markus Gerard**

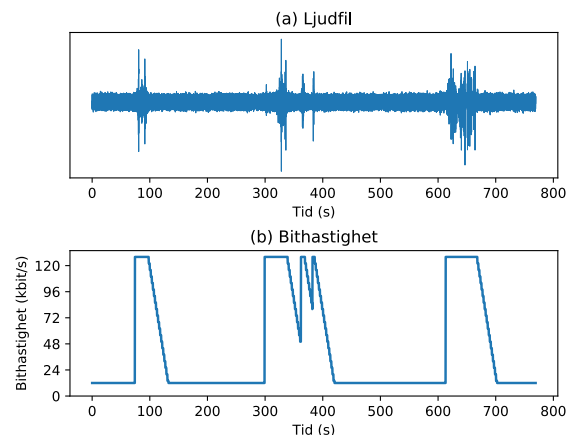
Omfattande forskning inom videokomprimering har lett till en märklig balans där ljudspåret ibland tar större plats än videospåret i övervakningsinspelningar. För att jämna ut detta behövs smarta tekniker även för ljudkomprimering. Det här arbetet presenterar ett förslag för att minska ljuddata med mer än 50 %.

Det är vanligt att majoriteten av materialet i en övervakningsinspelning är ointressant. Om man lyckas utnyttja det påståendet så finns det en möjlighet att effektivt minska datamängden genom att spara det ointressanta innehållet i låg kvalitet. Utmaningen ligger i att bestämma vad som är intressant och vad som är ointressant och att låta en dator ta de besluten.

För ljudinspelningar kan ointressant material vara tystnad eller brus, d.v.s. avsaknaden av aktivitet. Föreställ er en övervakningskamera inuti ett lager eller en butik som spelar in både ljud och bild dygnet runt. Under natten är aktiviteten låg och det enda som spelas in ljudmässigt är brus. Trots den låga aktiviteten sparas allt under natten med hög kvalitet, vilket resulterar i onödigt stora datafiler. Istället kan sådana långa sekvenser av tystnad komprimeras väldigt hårt eftersom dess resulterande kvalitet inte är viktig.

Det finns olika tillvägagångssätt för att känna igen aktivitet i ljud. En enkel lösning är att gå efter volymen i ljudet, om volymen överstiger bruset så förekommer det aktivitet. Nackdelen med den metoden är att ljudhändelser med samma eller lägre volym som bruset inte kommer att detekteras. En bättre lösning är att dela upp ljudet i frekvenser och sedan analysera ändringar i energi för olika frekvenser. De här energikomponenterna kan användas för att ta fram normalfördelningskurvor som kan modellera bakgrundsljudet. Ljud som avviker från bakgrundsmodellen klassificeras som aktivitet och kommer att komprimeras

med en hög bithastighet medan övrigt ljud komprimeras med en låg.



Jämfört med det traditionella sättet att spara allting med hög kvalitet oavsett innehåll så är det med den här metoden möjligt att minska datamängden med 55 % för miljöer där aktivitet förekommer 40 % av tiden. För miljöer med lägre aktivitet går det att minska ännu mer och vice versa.

Med ökande krav på mängden data att spara, bland annat från IoT-samhället som håller på att växa fram, är en lösning som detta en nödvändig utveckling för att hindra mängden data att bli ohållbart stor. Med en algoritm som denna kan man alltså välja att ha färre diskar med inspelat material, spara inspelningarna en längre tid, sätta upp fler kameror eller också öka kvaliteten på inspelningarna.