

EXAMENSARBETE Latent Space Interpolation for Music Transitions**STUDENT** Olle Rydén**HANDLEDARE** Ted Kronvall**EXAMINATOR** Maria Sandsten

Kan neurala nätverk skapa bättre övergångar mellan låtar?

POPULÄRVETENSKAPLIG SAMMANFATTNING

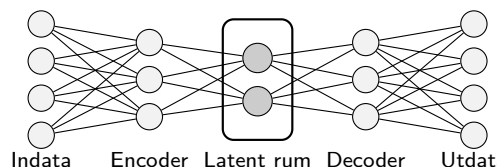
Olle Rydén

De flesta musikövergångar görs fortfarande med verktyg som beatmatchning och crossfades. I detta arbete undersöks om en neuralt tränad ljudkodare kan användas för att skapa övergångar direkt i ett komprimerat latent rum som skapas av neurala nätverk, och om resultatet kan mäta sig med traditionell signalbehandling.

När en DJ eller ett musikprogram byter från en låt till en annan vill man helst att bytet ska låta naturligt. I dag görs detta oftast med vanliga signalbehandlingsmetoder, där den enklaste automatiska varianten är att man låter den ena låten tona ut medan den andra tonar in. Det fungerar oftast bra, men metoden blandar bara två färdiga ljudsignaler ovanpå varandra.

I mitt examensarbete testar jag ett annat angreppssätt. I stället för att göra övergången direkt i ljudvågen kodas båda låtarna först in i ett *latent rum* med modellen EnCodec. Där representeras musiken som en tidsserie av kompakta vektorer som fortfarande bär viktig information om rytm, energi och tonart. Tanken är att man då kan interpolera mellan två låtar i denna representation och sedan avkoda resultatet tillbaka till ljud.

Metoden börjar med att välja passande segment ur två låtar. Dessa tempojusteras och radas upp så att taktslag och fraser stämmer överens. Därefter kodas segmenten till en kompakt representation, i det så kallade latent rummet. I detta latent rum testas flera interpolationsmetoder, bland annat linjär interpolering och sfärisk interpolering, så kallad SLERP. Resultatet avkodas sedan tillbaka till ljud och jämförs med en vanlig equal-power crossfade, alltså den typ av övergång som ofta används i praktiken.



Utvärderingen består av både lyssningstester och objektiva mått. I lyssningstesterna fick 14 personer jämföra 10 övergångar i det latent rummet med vanliga crossfades. Resultaten visade ingen tydlig vinnare: den latent metoden föredrogs ungefär lika ofta som vanliga crossfades, och många svar var "ingen preferens".

Samtidigt visar analyserna att latent interpolering inte beter sig exakt som en vanlig mix. Övergångarna blir ofta något mjukare och mindre detaljerade, särskilt i transienter och högfrekvent innehåll. Diagnostiken tyder också på att den latent representationen fångar musikalisk struktur, till exempel rytmisk periodicitet och tonhöjds-känslighet.

Slutsatsen är att latent interpolering inte ersätter den klassiska crossfaden, men att den är ett lovande alternativ. På sikt kan tekniken användas för smartare musikövergångar i automatiserade DJ-system, för spellistor och i musikverktyg.