

Computer herkent én noteert muziek

De meeste mensen herkennen *Yesterday* ook nog wel als het orkest het ritme tijdens het spelen wat verandert. Mensen kunnen kennelijk kleine afwijkingen van het verwachte patroon van toonhoogten, duur en positie van de noten interpreteren. Een computer daarentegen heeft er moeite mee. Het oplossen van dat soort herkenningproblemen, ook wel AI-complete problemen genoemd, vereist ‘menselijke intelligentie’. Toch zou het handig zijn als een computer van een willekeurig muziekstuk een transcriptie zou kunnen maken die door mensen te lezen en te begrijpen is. De meeste muziek is immers niet beschreven.

Ali Taylan Cemgil dacht tijdens zijn studie computerwetenschappen in Istanbul dat hij dat probleem wel in een half jaar op kon lossen. Inmiddels zijn we twaalf jaar verder en promoveert hij op 14 september aan de Radboud Universiteit Nijmegen. Hij is een heel eind, maar af is het nooit. Het kan altijd beter. “Een pianist die een stuk van bladmuziek speelt, zet die notatie om in geluidsgolven die wij horen. Dat is het omgekeerde van wat ik wil. Dus als ik daarvoor een model kan bedenken, hoef ik dat alleen nog maar te inverteren en dan kan ik dat gebruiken om een computer automatisch muziek te laten herkennen.” Voor het inverteren van die modellen bestaat een mathematisch goed doorwrocht model, gebaseerd op het theorema van Bayes.

Meest waarschijnlijk

Het klinkt eenvoudig, maar dat is het toch niet. Het lijkt op het volgende probleem: een excelfile rekent bijvoorbeeld het gemiddelde uit van drie getallen. Stel nu dat je alleen het eindresultaat kent en weet hoe je daaraan gekomen bent (door de functie ‘het gemiddelde nemen’) en je wilt daaruit de uitgangsggegevens bepalen. Dat leidt dan niet tot één, maar tot vele mogelijke antwoorden. Cemgil: “Wat je dan probeert, is door a priori kennis over het model toe te voegen, vast te stellen wat de meest waarschijnlijke uitgangsggegevens zijn geweest.”

Het afleiden van het gezochte model gaat in twee stappen. Het geluidssignaal bevat erg veel informatie. Daar kan een signaal uit afgeleid worden dat overeenkomt met een soort muziekrol zoals die in straatorgels gebruikt wordt. Dat levert een digitale representatie, een rijtje parameters: welke toets (toonhoogte) wordt wanneer ingedrukt en wanneer losgelaten (ritme en tempo). Een dergelijk file heet een midi-file (van *musical instrument digital interface*) en wordt gebruikt om bijvoorbeeld synthesizers aan te sturen. Het probleem is nu teruggebracht tot het relateren van dit midifile tot bijvoorbeeld onze westerse muzieknotatie met notenbalken. Daarvoor worden drempels aangebracht: duurt een bepaalde aanslag van een toon zo lang, dan komt dat overeen met een zestiende noot, bijvoorbeeld. Maar niet alle zestiende noten duren even lang, dus de computer moet met een marge werken.

Tempowisselingen

En om het echt lastig te maken zou die marge ook nog moeten kunnen veranderen gedurende de uitvoering. Cemgil: “Klassiek geschoolde pianisten spelen altijd met versnellingen en vertragingen ten opzichte van het basistempo. Amateurmuzikanten kun

je herkennen aan het feit dat ze naarmate het stuk vordert steeds sneller gaan spelen. Dat zou je een computer dus moeten vertellen, dat is het soort a priori kennis dat ik bedoel. Je moet de relatieve duur van een noot bepalen ten opzichte van het tempo waarin gespeeld wordt. Daarnaast kun je, als je weet dat de muziek op een piano gespeeld wordt, uit de fysische eigenschappen van het instrument gegevens afleiden omtrent het geluidssignaal. Ook die informatie rekenen we tot de a priori kennis. We gebruiken realistische aannames gebaseerd op kennis uit een groot aantal gebieden –cognitieve wetenschap, akoestiek, muzikwetenschap en computerwetenschap– om ons model te construeren.”

Het inverteren van dat model levert dan de meest waarschijnlijke oplossing voor de oorspronkelijke vraag, namelijk welke muzieknotatie hoort bij een bepaalde uitvoering. Die oplossing blijkt in de praktijk meestal heel goed te voldoen. Cemgil: “In Nijmegen in de vakgroep medische en biologische fysica werken we voornamelijk aan erg theoretische onderwerpen op dit gebied. Op een conferentie hadden we bij onze stand een piano staan waar mensen iets op konden spelen. De computer toverde dan de bijbehorende muziek in onze westerse notatie met notenbalken tevoorschijn. Dat was een mooie demonstratie van onze erg ingewikkelde modellen die op Bayesiaanse statistiek gebaseerd zijn. Deze zelfde technieken kun je in plaats van voor tempoveranderingen in muziek natuurlijk ook gebruiken om fluctuaties in andere variabelen te verklaren, in de beurskoers bijvoorbeeld.”

Toepassingen

Voor de automatische muziekherkenning zijn een aantal toepassingen te bedenken. Het genereren van bladmuziek door improviserende muzikanten is er maar één. Je kunt ook denken aan een systeem dat solisten begeleidt en zich aanpast aan het actuele tempo in plaats van dat de solist de begeleiding moet volgen. Of een muziektretrievalsysteem: je neuriet een paar maten en de computer zoekt voor je uit welk liedje maar niet uit je hoofd wil verdwijnen. Bovendien levert een model zoals Taylan Cemgil gebouwd heeft ook veel inzicht in hoe mensen muziek maken. En dat interesseert Cemgil, zelf een getalenteerd musicus die ooit zijn studie bijeenverdiende als muzikant, in hoge mate. “Ik hoop me in toekomst ook met dit soort computeronderzoek bezig te kunnen houden en het zou extra leuk zijn als het onderwerp nog steeds muziek is.”