

MATEMATICKÉ METÓDY PRI TRIEDENI HUDOBNÝCH ÚSEKOV

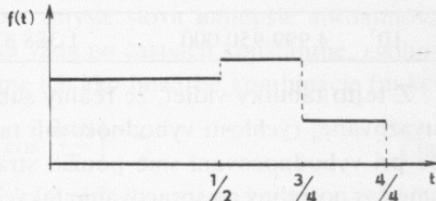
IGOR BALLO

Ako v každej vednej disciplíne, spojenej s rozsiahlym faktografickým materiáлом, tak aj v muzikológii, je veľmi dôležitá otázka triedenia, porovnávania, prípadne katalogizovania hudobných skladieb. Ako inde, tak aj tu, len roztriedením do skupín podľa vhodne definovanej podobnosti možno vytvoriť systém, podľa ktorého by sme sa mohli vo faktografickom materiáli lepšie vyznať, orientovať a vytvoriť tak podmienky pre odhalenie niektorých hlbších príbuzností.

Pre zisťovanie podobnosti a orientáciu v skúmaných hudobných materiáloch, nie je vždy potrebné a aj možné, v prvom štádiu skúmať celú skladbu. Podľa doterajších skúseností ako základ pre výskum sa vyberá hudobná téma alebo časti menších skladobných útvarov, pričom za najdôležitejšie informácie sa považuje melódia a rytmus. Týmto výberom vzniká tzv. hudobný úsek, ktorý definujeme ako jednohlasný melodickoritmický priebeh, zapisaný notovým písmom. Dôležitú úlohu má ohraničenie dĺžky hudobného úseku, ktorá určuje význam a zmysel z hľadiska skúmania podobnosti. Inými slovami môže to byť téma, motív, v ľudovej piesni melodický riadok a pod.

Pri posudzovaní podobnosti hudobných úsekov začala sa pocíťovať potreba obmedziť subjektívny pravok a zaviesť podľa možnosti matematicky formulovanú mieru podobnosti. V tomto smere bolo už viac pokusov, ktoré všetky vychádzali z vhodne zavedenej definícii vzdialosti dvoch prvkov súboru hudobných úsekov.

Základným pojmom pri definovaní tejto vzdialenosťi, ktorý treba zaviesť, je tzv. časová funkcia. Ide o vhodné matematické zobrazenie hudobných úsekov, najčastejšie v podobe funkčnej závislosti výšky noty (tónu) od jej trvania. Dostaneme tak po častiach konštantnú funkciu, graficky znázornenú napr. na obr. 1, kde sme vhodne zvolili mierku pre nezávisle premennú (čas) aj pre výšku (závisle premennú).



Obr. 1.

Definovať vzdialenosť dvoch takýchto časových funkcií a tým aj mieru podobnosti dvoch prvkov zo súboru hudobných úsekov, možno rôznym spôsobom. Možno ju napr. definovať [1] koeficientom podobnosti S_{ij} dvoch časových funkcií $f_i(t), f_j(t)$ ($i, j = 1, 2, \dots, n$)

$$S_{ij} = 1 - \frac{1}{F \cdot T} \sqrt{\int_0^T [f_i(t) - f_j(t)]^2 dt} \quad (1)$$

Z rovnosti (1) je zrejmé, že skôr ako vypočítame S_{ij} , musíme pretransformovať čas pomocou vhodnej mierky obidvoch časových funkcií tak, aby boli definované pre $t \in \langle 0, T \rangle$. Aby rovnosť (1) bola bezrozmerná, treba do nej zaviesť ešte normalizačný faktor F , ktorý napr. môžeme zvoliť ako $\max |f_i(t) - f_j(t)|$, alebo $\max (\max f_i, \max f_j) - \min (\min f_i, \min f_j)$. Tako zavedený koeficient bude z intervalu $\langle 0, 1 \rangle$ a pre identitné časové funkcie sa bude rovnať jednej (100 % podobnosť = identita).

Takýto spôsob kvantitatívneho hodnotenia podobnosti je veľmi jednoduchý a názorný. Prakticky sa však dá použiť len pri spracúvaní menších súborov hudobných úsekov, pretože počet koeficientov podobnosti sa rovná $n(n-1)/2$ a s počtom členov súboru n veľmi rýchlo narastá. Ako sa ľahko môžeme presvedčiť, pre celkom reálne súbory by sme dostali taký veľký počet týchto koeficientov, že ani pomocou modernej výpočtovej techniky by sme ich neboli schopní v reálnom čase všetky vypočítať. Okrem toho ich veľký počet ich znova robí pre nás bezcennými, pretože sa v nich nebudeme môcť orientovať. Konkrétnu situáciu ilustruje tab. 1.

Tabuľka 1

Počet prvkov súboru n	Počet koeficientov podobnosti S_{ij}	Čas, potrebný na vyčislenie všetkých koeficientov podobnosti, ak vyčislenie jedného trvá:	1 s	0,1 s	0,01 s
10^2	4 950	1 h 22,5 min		8,25 min	49,5 s
10^3	499 500	138 h 45 min		13 h 52,5 min	1 h 23,25 min
10^4	49 995 000	13 887 h 30 min		1 388 h 45 min	138 h 52,5 min
10^5	4 999 950 000	1 388 875 h		138 887 h 30 min	13 888 h 52,5 min

Z tejto tabuľky vidieť, že reálny súbor so 100 000 členmi by sme pri najvyššej uvažovanej rýchlosťi vyhodnocovali takmer dva roky. Vyplýva to zo skutočnosti, že pri vyhodnocovaní smie použili strategiu porovnávania *každý s každým*. Aby sme čas potrebný na spracovanie takýchto súborov, skrátili na reálne uskutočnitelnú hodnotu, musíme zmeniť práve základnú strategiu. Musíme použiť metódu,

která každý prvok súboru spracuje samostatne a podľa tohto jedného spracovania ho zaradí do niektoréj z tried podobných hudobných úsekov. V tomto prípade spracovanie celého súboru vyžaduje len n základných výpočtov. Ak by takýto základný počet pre jeden prvok trval napr. 10 sekúnd, potom spracovanie súboru so 100 000 prvkami by trvalo asi 278 hodín, čo je necelých 12 dní. Touto metódou, ktorú nazveme *globálnou*, môžeme celý súbor rozdeliť do viacerých tried, každú s menším počtom prvkov. V rámci týchto tried môžeme použiť skôr spomínanú metódu, pracujúcu strategiou *každý s každým*, ktorú nazveme *lokál-nou* metódou.

Globálnu metódu môžeme založiť napr. na známom princípe rozkladu na základné elementy. Tento princíp sa v prírodných a technických vedách už často veľmi dobre osvedčil a začína s úspechom prenikať aj do spoločenských vied. V optike to bol napr. princíp rozkladu svetla pomocou troch základných farieb, ktorým sa získala možnosť kvantitatívne charakterizovať farebné odtiene. Podobne v matematike rozvoj funkcií pomocou rôznych nezávislých sústav, tzv. súradných alebo bázových funkcií, môžeme chápať ako takýto princíp rozkladu na základné elementy.

Pri globálnom triedení hudobných úsekov sme použili práve tento princíp. Podľa neho sme príslušnú časovú funkciu $f(t)$ rozvinuli do radu

$$f(t) = \sum_{k=0}^n [\alpha_k \varphi_k(t) + \beta_k \psi_k(t)] + R_n(t) \quad (2)$$

kde φ_k , ψ_k sú nezávislé funkcie, definované pre $t \in \langle 0, T \rangle$. Ak zvyšok R_n je v istom zmysle dostatočne malý, potom funkciu f , pri vopred zadaných funkciach φ_k , ψ_k , možno charakterizovať číselnými koeficientmi α_k , β_k . Udávajú akýsi obsah jednotlivých základných súradných funkcií φ_k , ψ_k v časovej funkcií f . Dá sa ukázať — a v ďalšom to ukážeme len na príklade — že pre dve časové funkcie, v istom zmysle slova blízke, sa číselne najväčšie z koeficientov α_k , β_k pri obidvoch funkciách len málo líšia. Je zrejmé, že podľa tejto vlastnosti môžeme už vykonať triedenie časových funkcií, a tým aj hudobných úsekov.

Pre úspešné použitie tohto spôsobu triedenia sme predpokladali malosť zvyšku R_n . Dosiahneme ju, ak použijeme sústavu súradných funkcií φ_k , ψ_k , pomocou ktorej pri pevnom n budeme môcť v istom zmysle slova najlepšie approximovať časovú funkciu f . Pretože časové funkcie sú vždy po častiach konštantné, zvolíme aj φ_k , ψ_k ako takéto funkcie. Skonštruujujeme ich ako lineárne kombinácie funkcií

$$\varphi_i(t) = \operatorname{sgn} \left[\cos \frac{2\pi l t}{T} \right] \quad (3a)$$

$$\psi_i(t) = \operatorname{sgn} \left[\sin \frac{2\pi l t}{T} \right] \quad (3b)$$

$$\varphi_0(t) = 1; \quad \psi_0(t) = 0 \quad (3c)$$

$$l = 1, 2, \dots t \in \langle 0, T \rangle$$

Súradnice funkcie $\varphi_k(t)$, $\psi_k(t)$ potom budú:

$$\varphi_k(t) = \sum_{l=0}^k (c_{lk}\varphi_l^* + d_{lk}\psi_l^*) \quad (4a)$$

$$\psi_k(t) = \sum_{l=0}^k (e_{lk}\varphi_l^* + f_{lk}\psi_l^*) \quad (4b)$$

s príavnou požiadavkou ich vzájomnej ortonormovanosti. Znamená to, že pre ne vyžadujeme platnosť vzťahov

$$\int_0^T \varphi_k(t) \varphi_j(t) dt = \begin{cases} 0 & \dots j \neq k \\ 1 & \dots j = k \end{cases} \quad (5a)$$

$$\int_0^T \varphi_k(t) \psi_j(t) dt = 0 \quad (5b)$$

$$\int_0^T \psi_k(t) \varphi_j(t) dt = 0 \quad (5c)$$

$$\int_0^T \psi_k(t) \psi_j(t) dt = \begin{cases} 0 & \dots j \neq k \\ 1 & \dots j = k \end{cases} \quad (5d)$$

$$k = 0, 1, 2, \dots; \quad j = 0, 1, 2, \dots, k$$

Takto sústava (5) tvorí $4(k+1)$ rovníc pre $4(k+1)$ neznámych koeficientov c_{lk} , d_{lk} , e_{lk} , f_{lk} , vystupujúcich v rovniciach (4). Na jej základe môžeme ďalej jednoducho vyrátať číselné koeficienty α_k , β_k , ktoré budú:

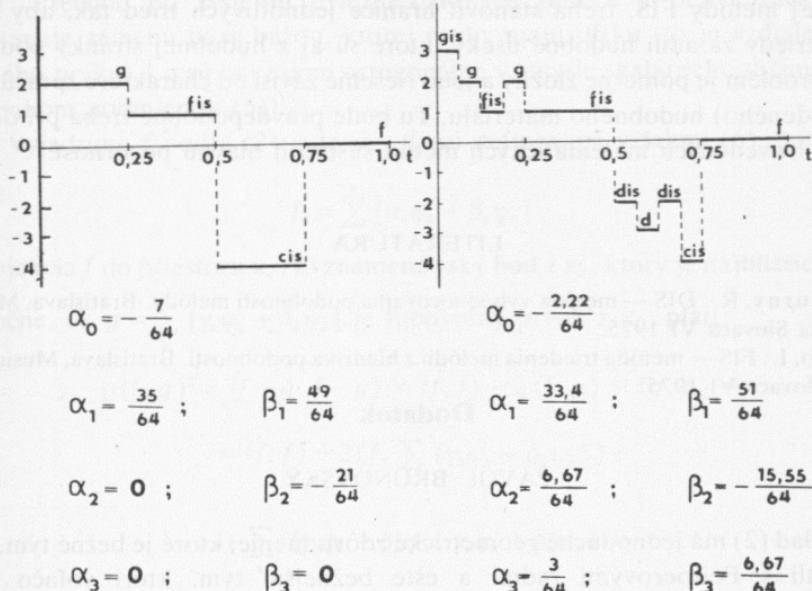
$$\alpha_k = \int_0^T f(t) \varphi_k(t) dt \quad (6a)$$

$$\beta_k = \int_0^T f(t) \psi_k(t) dt \quad (6b)$$

Pre ilustráciu uvedieme jednoduchý príklad použitia tejto globálnej metódy, ktorú v pracovných materiáloch nazývame metódou FIS. Grafické znázornenie dvoch jednoduchých časových funkcií s príslušnými koeficientmi α_k , β_k (pre $k = 0, 1, 2, 3$) je na obr. 2. Pri jednotlivých konštantných úsekoch obidvoch funkcií sú pre názornosť pripísané zodpovedajúce výšky nôt.

Z tohto obrázku a príslušných koeficientov je zrejmé, že obidve časové funkcie sú charakteristické tým, že sú pri nich relatívne najväčšie koeficienty α_1 , β_1 a že sú navzájom blízke. Koeficient β_2 má zrejme len menší význam, kým ostatné koeficienty sú pre charakterizovanie blízkosti obidvoch funkcií, t. j. podobnosti

príslušných hudobných úsekov, bezvýznamné. Podľa týchto zásad by sme mohli k týmto dvom časovým funkciám priradiť aj ďalšie, a tým vytvoriť triedu, charakterizovanú veľkými koeficientmi α_1 , β_1 , menším koeficientom β_2 a málo významnými ostatnými koeficientmi. Prirodzene, že pri konkrétnnej realizácii by sme hranice triedy museli presne definovať.



Obr. 2.

Tieto všeobecné úvahy našli už aj praktické uplatnenie. V spolupráci s hudobným oddelením Slovenského národného múzea v Bratislave s niektorými pracovníkmi Ústavu mechaniky strojov SAV sa vypracovali viaceré programy pre číslicový počítač GIER (programovací jazyk GIER Algol 4), ktoré zabezpečujú:

Program EIS: vstup údajov o hudobnom úseku do počítača pomocou osobitného alfanumerického kódu (M-ALMA-72*), kontrola syntaktických pravidiel notového písma, výpočet a vytlačenie normalizovanej časovej funkcie.

Program FIS: z vypočítanej časovej funkcie vypočíta koeficienty α_k , β_k . Ďalej ich roztriedi podľa špeciálneho algoritmu, ktorý sme tu neopísali.

Program DIS: vychádza tiež z časových funkcií, pre ktoré počíta koeficient podobnosti S_{ij} [1].

Všetky tieto programy zostavil a odladil Ing. R. Chmúrny, CSc. Pomocou nich

* Je to programovací jazyk, blízky notovému písma, podobne ako jazyk ALGOL 60 blízky matematickému zápisu výrazov.

v Slovenskom národnom múzeu (dr. L. Ballová) spracovali viaceré súbory hudobných úsekov. Na týchto testovacích súboroch sa skúma miera citlivosti a účinnosti zvolených matematických metód predovšetkým z hudobného hľadiska. Dosiahnuté výsledky, hoci zodpovedajú základným predpokladom a požiadavkám, predsa treba pokiaľ za predbežné. Dôvod je v tom, že pre úspešné použitie globálnej metódy FIS, treba stanoviť hranice jednotlivých tried tak, aby sa do jednej triedy zaradili hudobné úseky, ktoré sú aj z hudobnej stránky podobné. Tento problém je pomerne zložitý a jeho riešenie závisí od charakteru spracúvaného (triedeneho) hudobného materiálu. Tu bude pravdepodobne treba pri ďalšom rozvíjani uvedených matematických metód sústrediť hlavnú pozornosť.

LITERATÚRA

1. Chmúrny, R.: DIS — metóda vyhodnocovania podobnosti melódii. Bratislava, Musicologica Slovaca VI 1975.
2. Ballo, I.: FIS — metóda triedenia melódii z hľadiska podobnosti. Bratislava, Musicologica Slovaca VI 1975.

Dodatok

PAVOL BRUNOVSKÝ

Rozklad (2) má jednoduché geometrické zdôvodnenie, ktoré je bežné tým, ktorí sa stretli s Fourierovými radmi a ešte bežnejšie tým, ktorí voľačo vedia z funkcionálnej analýzy.

Za priestor hudobných úsekov χ vezmeme množinu po čiastkach konštantných funkcií na intervale $\langle 0, T \rangle$. Je to lineárny priestor, čo znamená, že spolu s dvoma funkiami obsahuje aj ich ľubovoľnú lineárnu kombináciu. V tomto priestore zavedieme skalárny súčin predpisom

$$\langle f, g \rangle = \int_0^T f(t) \cdot g(t) dt \quad (7)$$

pomocou ktorého definujeme vzdialenosť dvoch funkcií podobne ako vzdialenosť bodov v euklidovskom priestore

$$\varrho(f, g) = \langle f - g, f - g \rangle^{1/2} = \left[\int_0^T (f(t) - g(t))^2 dt \right]^{1/2}$$

(koeficient podobnosti, definovaný v (1), sa potom rovná $S_{ij} = (FT)^{-1} \varrho(f_i, f_j)$). Možno dokázať, že takto definovaná vzdialenosť má vlastnosti, ktoré od vzdialenosť obvykle vyžadujeme: $(\varrho(f, g) \geq 0, \varrho(f, g) = 0 \Leftrightarrow f = g, \varrho(f, g) = \varrho(g, f) \text{ a } \varrho(f, g) + \varrho(g, h) \geq \varrho(f, h) — \text{trojuholníková nerovnosť})$.

Označme κ_n lineárny obal funkcií $\varphi_i^*, \chi_i^*, j = 1, \dots, n$, čo je najmenšia množina, uzavretá vzhľadom na lineárne operácie, obsahujúca $\varphi_i^*, \psi_i^*, j = 1, \dots, n$ (geometricky je to $2n$ -rozmerná rovina, obsahujúca vektory φ_i^*, ψ_i^*). Vektory $\varphi_i^*, \psi_i^*, i = 1, \dots, n$ tvoria bázu lineárneho priestoru κ_n , t. j. sú lineárne nezávislé a každý prvok z κ_n sa dá jediným spôsobom vyjadriť ako lineárna kombinácia $\varphi_i^*, \psi_i^*, j = 1, \dots, n$. Prechod od systému hviezdičkovaných funkcií k nehviezdičkovaným predstavuje zámennu bázy bázou, ktorej prvky majú dĺžku (to je vzdialenosť od nulového prvku) 1 a sú navzájom ortogonálne v zmysle skalárneho súčinu (7) (to je obsahom podmienok (5)).

Ak vyjadríme f podľa (2), kde α_k, β_k sú definované vzťahmi (6), potom

$$f_0 = \sum_{k=0}^n [\alpha_k \varphi_k + \beta_k \psi_k]$$

je projekcia f do priestoru κ_n , to znamená taký bod z κ_n , ktorý je najbližšie bodu f .

Skutočne, ak $g = \sum_k [\gamma_k \varphi_k + \delta_k \psi_k]$ je ľubovoľný prvok z κ_n , platí:

$$\begin{aligned} \varrho(f, g)^2 &= \langle f - g, f - g \rangle = \langle f, f \rangle - 2 \langle f, g \rangle + \langle g, g \rangle = \\ &= \langle f, f \rangle - 2 \langle f, \sum_k (\gamma_k \varphi_k + \delta_k \psi_k) \rangle + \\ &\quad + \sum_k \langle \gamma_k \varphi_k + \delta_k \psi_k, \gamma_k \varphi_k + \delta_k \psi_k \rangle = \\ &= \langle f, f \rangle - 2 \sum_k [\gamma_k \langle f, \varphi_k \rangle + \delta_k \langle f, \psi_k \rangle] - \\ &\quad - \sum_{k \neq j} [\gamma_k \gamma_j \langle \varphi_k, \varphi_j \rangle + \gamma_k \delta_j \langle \varphi_k, \psi_j \rangle + \\ &\quad + \gamma_j \delta_k \langle \varphi_j, \psi_k \rangle + \delta_k \delta_j \langle \psi_j, \psi_k \rangle]. \end{aligned}$$

Vzhľadom na vzťahy (5) z toho dostaneme:

$$\begin{aligned} \varrho(f, g)^2 &= \langle f, f \rangle - 2 \sum_k (\gamma_k \alpha_k + \delta_k \beta_k) + \sum_k (\gamma_k^2 + \delta_k^2) = \\ &= \langle f, f \rangle + \sum_k [(\gamma_k - \alpha_k)^2 + (\delta_k - \beta_k)^2] - \sum_k (\alpha_k^2 + \beta_k^2). \end{aligned}$$

Je zrejmé, že $\varrho(f, g)^2$ nadobúda minimum práve vtedy, keď $\gamma_k = \alpha_k, \delta_k = \beta_k$, t. j. ak $g = f_0$.

Ak by sme teda napr. za kombinácie hudobných úsekov φ_k^*, ψ_k^* považovali všetky hudobné úseky, ktoré dostaneme lineárnu kombináciou úsekov φ_k^*, ψ_k^* , bola by f_0 kombinácia úsekov φ_k^*, ψ_k^* najbližšia k f .