

АНОТАЦІЯ

Комаров О.В. Моделі та методи динамічного музичного синтезу на основі генетичного підходу. – Кваліфікаційна наукова праця на правах рукопису.

Дисертація на здобуття наукового ступеня доктора філософії за спеціальністю 122 - Комп'ютерні науки. – Державний університет «Одеська політехніка» МОНУ, Україна, Одеса, 2021.

У *вступі* обґрунтовано актуальність вирішення завдання розробки методів та моделей динамічного музичного синтезу з високим рівнем гармонійності в умовах обмежених обчислювальних можливостей. Динамічний музичний синтез (ДМС) – це автоматизований процес утворення музичних композицій із заздалегідь створених мелодійних фрагментів за рахунок їх гармонізації. Показано, що ДМС потребує значної обчислювальної потужності та, як правило, вимагає узгодженості із зовнішніми керуючими процесами, наприклад, діями користувача ігрового додатку або процесом створенням медійного контенту в реальному часі. Визначено об'єкт, предмет, задачі і методи дослідження; показано зв'язок з науковими програмами та планами; наведено наукову новизну та практичне значення одержаних результатів; висвітлено особистий внесок здобувача.

У *першому розділі* дисертаційної роботи проведено аналіз проблем створення систем динамічного музичного синтезу з високим рівнем гармонійності.

Суть задачі ДМС полягає у створенні композиції шляхом формування акомпанементу для заданої мелодії. При цьому, композиція представляється як послідовність сукупностей одночасно вжитих звуків (акордів) до кожного з яких входить один звук мелодії. Гармонізація виконується над множиною мелодій, написаних композитором заздалегідь.

Проаналізовано існуючі методи синтезу музичних композицій, їх переваги та недоліки. Показано, що побудова на їх основі систем динамічного

музичного синтезу (СДМС) ускладнена, так як вони синтезують музичні композиції, що не в повній мірі відповідають вимогам гармонійності, та/або не можуть задовольнити ці вимоги в режимі реального часу при обмеженій обчислювальній потужності. Сформульовано протиріччя, яке вирішується в дисертаційній роботі, – між вимогами до гармонійності та можливостями їх забезпечення при обмеженні часу та обчислювальною потужністю.

Проведено аналіз множини існуючих музичних систем для інтерактивних мультимедійних проєктів. Запропоновано виокремити такі, що є найбільш розповсюдженими та представляють особливий інтерес: діатонічну музичну систему та системи з регулярними ладами (рис. 1).



Рисунок 1 – Схема взаємозв'язку розповсюджених музичних систем.

Показано, що найбільш перспективним підходом для розробки методів ДМС є еволюційний підхід. Його застосування передбачає розробку моделей генотипних представлень для розглянутих музичних систем та метода генетичної гармонізації.

Генетичні методи для гармонізації дозволяють сформувати низку еквівалентних, різних за змістом рішень, що відрізняються високим рівнем гармонійності отриманих результатів, так як фітнес-функція генетичного методу $H(c)$ при створенні композиції C враховує поняття гармонійності та визначається за допомогою набору правил сполучення звуків, і має вигляд:

$$H(C) = 1 - \frac{1}{Q(L-1)} \sum_{i=1}^{|F|} \sum_{j=1}^{L-1} q(f_i) f_i(\mathbf{A}_j, \mathbf{A}_{j+1}) \quad (1)$$

де L – кількість акордів (сполучень одночасно вжитих звуків) композиції, Q – сума ваг всіх правил вживання акордів, f_i – i -те правило вживання акордів, $i = 1$,

... , $|F|$, $|F|$ – кількість правил вживання акордів, q – функція ваг правил вживання акордів, A_j – j -тий акорд композиції, $j = 1, \dots, L-1$.

Недоліком існуючих генетичних методів гармонізації є низька оперативність через надмірність області пошуку, на якій виконується пошук придатного рішення. Область пошуку відповідає простору параметрів генетичних представлень, і її розмір залежить від «вдалості» їх кодування. Як правило, для забезпечення пошуку по всій множині фенотипів, розробники генетичних представлень віддають перевагу повноті представлення та не задовольняють вимогу допустимості. Показано, що підвищення допустимості генетичного представлення музичної композиції шляхом врахування особливостей конкретної музичної системи, для якої виконується гармонізація, може бути резервом для зменшення області, та, відповідно, і часу пошуку.

Сформульовано мету дисертаційного дослідження та його задачі.

В *другому розділі* було розроблено консонантну акордну модель генотипного представлення музичної композиції для діатонічної музичної системи та генетичний метод гармонізації з її використанням. При розробці моделі для підвищення допустимості генотипного представлення запропоновано застосування *принципу придатної повноти*, що передбачає існування генотипів тільки для придатних рішень за рахунок підвищення структурної узагальненості представлення. Структурна узагальненість визначається рівнем абстрагування музичних концепцій параметрами представлення. Виділено три її рівні – звуковий, ладовий та акордний.

В рамках найвищого, акордного, рівня структурної узагальненості на основі дослідження структури акорду тризвуку діатонічної системи запропоновано виділити набір параметрів:

- функція $f \in \{1, 2, 3\}$ визначає початок терцового порядку, його ступінний склад і роль в послідовності акордів;
- обернення $i \in \{1, 2, 3\}$ визначає початок голосового порядку та ступінь найнижчого звуку акорду;
- розміщення $p \in \{1, 2\}$ визначає відносне положення середніх голосів;

– повтор $r \in \{1, 3\}$ визначає ступінь, що буде відтворено двома звуками.

На основі цього набору запропоновано консонантну акордну модель генотипного представлення музичних композицій $\mathbf{C}^G$ у вигляді послідовності образів акордів $\mathbf{A}_j^G = (m_j, f_j, i_j, p_j, r_j)$:

$$\mathbf{C}^G = \begin{bmatrix} m_1 & m_2 & \dots & m_L \\ f_1 & f_2 & \dots & f_L \\ i_1 & i_2 & \dots & i_L \\ p_1 & p_2 & \dots & p_L \\ r_1 & r_2 & \dots & r_L \end{bmatrix} \quad (2)$$

де m_j – звуки мелодії, $j = 1, \dots, L$, L – довжина мелодії.

Фенотипним представленням музичної композиції $\mathbf{C}$ для генетичного методу з фітнес-функцією (1) є матриця, що є послідовністю з L акордів $\mathbf{A}_j$:

$$\mathbf{C} = \{\mathbf{A}_j : \mathbf{A}_j = \{P_{kj} | k = 1, \dots, 4\}\} \quad (3)$$

де P_{kj} – значення висот окремих звуків.

Для декодування моделі (2) в фенотип (3) запропоновано функцію *pitch*, що відображає кожен j -тий стовпчик матриці $\mathbf{C}^G$ в акорд $\mathbf{A}_j$:

$$\mathbf{A}_j = \text{pitch}(\text{order}(\text{degrees}(m_j, f_j), i_j, p_j, r_j), m_j) \quad (4)$$

де *degrees* – функція, що утворює вектор ступенів, вживаних в акорді, розташованих в терцовому порядку, *order* – функція, що утворює вектор ступенів, відповідних звукам акорду, розташованих в голосовому порядку.

$$\text{degrees}(m, f) = (\delta^d : \delta^d = (7 + \delta(m) - 2(f - d)) \bmod 7)$$

де δ^d – ступінь ладу, $d \in \{1, 2, 3\}$, δ – функція відповідності між висотами та ступенями ладу, *mod* – операція цілочисельного ділення.

Функція *order*(Δ, i, p, r) впорядковує та розширює утворений функцією *degrees* вектор ступенів терцового порядку Δ , формуючи новий вектор ступенів голосового порядку $\Delta_P = (\delta_1, \delta_2, \delta_3, \delta_4)$:

$f \neq i$	$f = i = 2$	$f = i \neq 2$
$\delta_1 = \delta^f$ $\delta_{1+p} = \Delta \setminus \{\delta^f, \delta^i\}$ $\delta_{4-p} = \delta^r$ $\delta_4 = \delta^i$	$\delta_1 = \delta^f$ $\delta_2 = \delta^p$ $\delta_3 = \delta^{3-p}$ $\delta_4 = \delta^r$	$\delta_1 = \delta^f$ $\delta_{1+p} = \Delta \setminus \{\delta^f, \delta^2\}$ $\delta_{4-p} = \delta^2$ $\delta_4 = \delta^i$

Функція *pitch* перетворює вектор ступенів голосового порядку Δ_P на вектор висот A_j на основі значення висоти мелодії m :

$$pitch(\Delta_P, m) = \left\{ P_k : P_k = \begin{cases} m & k = 1 \\ \delta^{-1}(\delta_k, \omega(P_{k-1})) & P_k < P_{k-1} \\ \delta^{-1}(\delta_k, \omega(P_{k-1}) - 1) & P_k \geq P_{k-1} \end{cases} \right.$$

де δ^{-1} – функція відповідності ступеню ладу δ_k і октави ω_k звуку з його висотою, $k \in \{1, 2, 3, 4\}$.

Таким чином, отримано перший пункт наукової новизни: вперше запропоновано консонантну акордну модель генотипного представлення музичних композицій (2), особливістю якої відповідність принципу придатної повноти за рахунок використання акордного рівню структурної узагальненості, що дозволило зменшити область пошуку рішення в генетичних методах.

На основі моделі (2) запропоновано генетичний метод гармонізації, який передбачає наступні кроки:

1. Сформувати початкову популяцію з N осіб у відповідності з обраною моделлю представлення, наприклад, для діатонічної системи – C^G (2).

2. Встановити лічильник ітерацій зі значенням 0.

3. Для обчислення фітнес-функції $H(C)$ (1) декодувати популяцію в фенотипне представлення (3) з урахуванням відповідної моделі обраної музичної системи. Наприклад, у разі діатонічної – за допомогою функції (4).

4. Якщо значення фітнес-функції $H(C)$ (1) не менше наперед заданого порогу або значення лічильнику ітерацій більше наперед заданого порогу, закінчити роботу. В іншому випадку, збільшити лічильник ітерацій на 1 і перейти до наступного пункту.

5. Відібрати n найбільш гармонійних осіб популяції C^G .
6. Утворити нову популяцію розміру N одноточковим кросинговером між кожною парою із n відібраних гармонійних осіб.
7. Провести мутацію популяції, з шансом p для кожної її особи.
8. Повернутись до пункту 2.

Дослідження показали, що використання запропонованого генетичного методу гармонізації дозволяє в 4 рази скоротити кількість ітерацій, необхідних для досягнення рівня гармонійності 0.9 у порівнянні з методом без урахування моделі (2).

Запропонований метод може бути використаний для гармонізації з використанням інших моделей генотипного представлення та становить другий пункт наукової новизни.

В *третьому розділі* розроблено метод створення моделей генотипного представлення музичних композицій для довільної музичної системи та системи з регулярним ладом на основі використання нейромережевого автокодувальника для декодування параметрів акордів.

Нейромережевий автокодувальник – це багат шарова нейромережа симетричної архітектури з вузьким центральним шаром, що навчається формувати на центральному шарі образ вхідних даних меншої розмірності.

Запропонована модель генотипного представлення композиції C^E для довільної музичної системи та системи з регулярним ладом у вигляді послідовності акордів A_j^E :

$$C^E = \{A_j^E : A_j^E = \{m_j\} \cup \{a_{ej} : a_{ej} \in \{0, 1\}\}\} \quad (5)$$

де m_j – звуки мелодії, $j = 1, \dots, L$, L – довжина мелодії. a_{ej} – параметри акорду, $e = 1, \dots, E$, E – обраний розмір кодування акорду.

На основі сформованої вибірки акордів заданої музичної системи навчається автокодувальник з центральним шаром розміру E . Під час ініціалізації методу генетичної гармонізації створюється популяція особин представлення (5), де параметри a_{ej} заповнюються випадковими значеннями.

На етапі оцінки особини створена популяція декодується в фенотипне представлення (3).

Для довільної музичної системи розроблено метод створення моделей генотипного представлення музичних композицій шляхом навчання автокодувальника, що передбачає наступні кроки:

1. Перетворити кожен вектор висот звуків акорду $\mathbf{A}_s^\sigma = \{P_{ks}\}$ з вибірки на вектор інтервалів $\mathbf{A}_s^\Delta$:

$$\mathbf{A}_s^\Delta = \{\Delta P_{ks} : \Delta P_{ks} = P_{(k-1)s} - P_{ks}\}$$

де $k = 2, \dots, 4, s = 1, \dots, S, S$ – розмір вибірки.

2. Побудувати гістограму значень ΔP_{ks} . На основі наперед заданої допустимої частоти вживання інтервалів α , знайти найменший інтервал ΔP_m для якого $p(\Delta P_m) < \alpha$.

3. Вилучити вектори $\mathbf{A}_s^\Delta$, що містять інтервали $\Delta P_{ks} \geq \Delta P_m$.

4. Перетворити значення інтервалів на двійкові вектори довжини b , де b – натуральне число, що відповідає наступній нерівності:

$$2^{b-1} < \Delta P_m \leq 2^b$$

5. Конкатенувати двійкові вектори інтервалів одного акорду.

6. Провести навчання автокодувальника з центральним шаром розміру E з використанням отриманої множини векторів.

7. Сформувати декодер шляхом вилучення з автокодувальника вхідного та наступних шарів до центрального і заміною центрального шару на вхідний.

Під час обчислення фітнес-функції, двійкова частина кожного акорду вигляду (5) оброблюється декодером, розподіляється на окремі вектори, що проходять дебінаризацію в десяткові значення ΔP_{ks} та перетворюються на значення висот P_{ks} , відштовхуючись від значення P_{1a} , формуючи фенотипне представлення акорду.

Модель C^E може бути використана в генетичному методів гармонізації. Експериментально доведено, що використання запропонованої моделі з таким декодуванням дозволяє в 1,8 рази скоротити кількість ітерацій, необхідних для

досягнення рівня гармонійності 0.9 для довільної музичної системи у порівнянні з методом без урахування моделі C^E .

Метод створення моделей генотипного представлення композицій для музичних систем з регулярними ладами відрізняється від попереднього обчисленням інтервалів не значень висот, а значень відповідних їм номерів ступенів ладу. Для цього попередньо проводиться реконструкція ладу:

1. Побудувати гістограму розподілу вжитих в акордах A_s^σ вибірки висот P_{ks} по висотним класам, де висотний клас - це клас еквівалентності, до якого відносяться будь-які дві висоти P_ϕ і P_ψ , для яких справедлива рівність:

$$P_\psi = 2^n P_\phi, n \in \mathbb{Z}$$

2. На основі прийнятої допустимої частоти вживання висотних класів γ , вилучити такі висотні класи P^{Ci} , для яких $p(P^{Ci}) < \gamma$.

3. Встановити порядок на висотних класах, визначивши функцію:

$$\delta : P \rightarrow D, D \subset \mathbb{N}$$

де P – множина усіх можливих висот, D – множина номерів ступенів ладу.

Ладові інтервали обчислюються наступним чином:

$$P_{1s} \ominus P_{2s} = (\delta(P_{1s}) - \delta(P_{2s})) \bmod |D|$$

Пункти 2 і 3 методу для довільної музичної системи можна не проводити, оскільки b відповідає наступній нерівності:

$$2^{b-1} < |D| \leq 2^b$$

Оскільки значення b в цьому методі менше, ніж в методі для довільної музичної системи, це дозволяє зменшити розмір кодування E при навчанні автокодувальника, зменшуючи простір ознак представлення акорду.

Під час обчислення фітнес-функції, двійкова частина кожного акорду вигляду (5) оброблюється декодером, розподіляється на окремі вектори, що проходять дебінаризацию в десяткові значення, а згодом перетворюються на значення ступенів акорду, розташованих в голосовому порядку, утворюючи множину Δ_P . Вектор висот фенотипу утворюється функцією (4), обчисленою на Δ_P та P_{Is} .

Експериментально доведено, що використання запропонованої моделі з таким декодуванням в розробленому методі гармонізації дозволяє в 3,5 рази скоротити кількість ітерацій, необхідних для досягнення рівня гармонійності 0.9 в системах з регулярними ладами у порівнянні з методом без урахування моделі C^E .

Таким чином, отримано третій пункт наукової новизни: вперше запропоновано метод створення моделей генотипного представлення музичних композицій для довільної музичної системи та систем з регулярними ладами, особливістю яких є відповідність принципу придатної повноти за рахунок використання нейромережевого автокодувальника для досягнення акордного рівню структурної узагальненості, що дозволило зменшити область пошуку рішення в генетичних методах та тривалість гармонізації.

В *четвертому розділі* розроблено структуру системи динамічного музичного синтезу (рис. 2), запропоновано модель оцінки необхідної обчислювальної потужності СДМС на основі генетичного підходу.

Представлена система пов'язує стани системи за допомогою автомату, що показує можливу зміну одного стану на інший. Зовнішній процес, з яким узгоджується робота СДМС, виконує роль вхідного сигналу для цього автомату. Репертуар містить музичні теми, що співвіднесені зі станами автомату. Генератор областей переходу утворює мелодію для гармонізації шляхом об'єднанням фрагменту поточної теми з наступним фрагментом поточної і початковими фрагментами тем суміжних станів. Залежно від музичної системи, в якій утворено мелодію, конфігуратор генетичного методу підбирає правила гармонійності та відповідні моделі представлення – консонантну акордну модель (2) або довільна кількість моделей (5), створених з використанням автокодувальника. Генетичний метод гармонізує композиції для усіх можливих сценаріїв зміни поточного стану, додаючи їх до відповідного банку у вигляді фенотипів (3).

Коли надходить сигнал з інформацією про поточний стан, система відтворює один із попередньо гармонізованих сценаріїв, що відповідає

поточному стану, та починає гармонізацію сценаріїв після оновлення стану автомату.

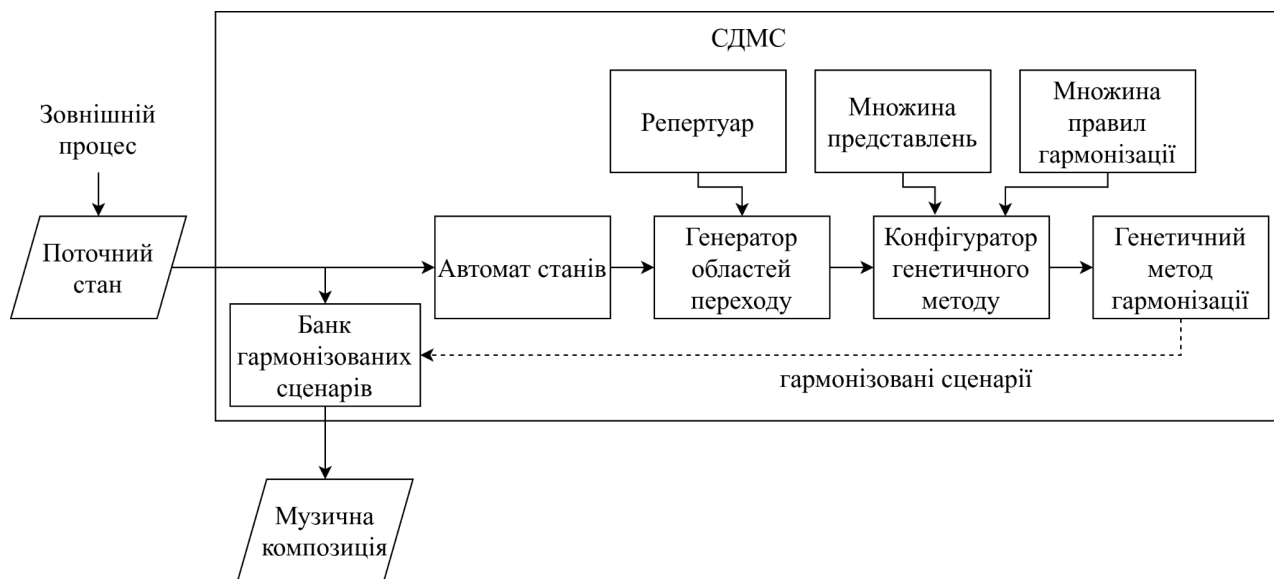


Рисунок 2 – Структурно-функціональна схема СДМС.

На основі залежностей між максимальною допустимою тривалістю гармонізації, кількістю акордів, що гармонізуються, і тактів мелодій запропоновано два різних режими роботи СДМС: швидкісний та якісний.

Запропоновано модель розрахунку обчислювальних вимог розроблюваної СДМС. Необхідна обчислювальна потужність M для гармонізації репертуару сценаріїв за t_{min} секунд складає:

$$M = \sum_{\omega}^{\Omega} \frac{IN(L_{\omega} - 1)\lambda_{\omega}}{t_{min}} \quad (6)$$

де Ω – найбільша множина областей переходу для поточного стану системи, ω – область переходу конкретного сценарію, I – кількість ітерацій генетичного алгоритму, N – розмір популяції, L_{ω} – кількість мелодичних звуків області переходу ω , λ_{ω} – число операцій, необхідне для обчислення гармонійності однієї пари акордів на представленні, що асоційована з областю переходу ω .

Модель (6) дозволила аналітично визначати необхідну обчислювальну потужність для роботи СДМС в заданих часових обмеженнях на основі генетичного методу з використанням розроблених генотипних представлень

акордного рівня структурної узагальненості (2) і (5) та моделі представлення меншого – ладового – рівня (табл.1 – табл.3).

Таблиця 1 – Характеристики гармонійності та обчислювальних потужностей, потрібних для гармонізації одного сценарію в швидкісному режимі роботи.

H	Модель ладового рівня		Модель $C^E(5)$ для системи з рег. ладом		Модель $C^G(2)$	
	I	$M^I, 10^6$ FLOPS	I	$M^I, 10^6$ FLOPS	I	$M^I, 10^6$ FLOPS
0.61	1	0.208	1	0.065	1	0.049
0.7	3	0.697	1	0.065	1	0.049
0.82	6	1.659	3	0.225	1	0.049
0.91	13	3.647	5	0.388	4	0.212
0.94	31	8.972	8	0.635	5	0.269

Таблиця 2 – Характеристики гармонійності та обчислювальних потужностей, потрібних для гармонізації одного сценарію в якісному режимі роботи.

H	Модель ладового рівня		Модель $C^E(5)$ для системи з рег. ладом		Модель $C^G(2)$	
	I	$M^I, 10^6$ FLOPS	I	$M^I, 10^6$ FLOPS	I	$M^I, 10^6$ FLOPS
0.61	3	2.999	1	0.402	1	0.13
0.7	6	7.108	3	1.381	1	0.13
0.82	9	11.772	6	3.156	1	0.13
0.91	14	19.787	9	4.892	6	0.874
0.97	35	52.435	28	15.959	20	3.045

Порівняння тривалостей гармонізації було виконано на апаратній платформі з характеристиками: Intel(R) Xeon(R) CPU @ 2.20GHz (1 core, 2 threads, 32 bit, cache 512 kb), 13 GB RAM. Випробування було проведено в двох режимах для двох рівнів гармонійності.

Таблиця 3 – Час, потрібний для гармонізації одного сценарію в різних режимах.

H	Швидкісний режим			Якісний режим		
	Модель ладового рівня	Модель $C^E(5)$ з рег. ладом	Модель $C^G(2)$	Модель ладового рівня	Модель $C^E(5)$ з рег. ладом	Модель $C^G(2)$
0.7	414 мс	93 мс	34 мс	13.6 с	3.9 с	269 мс
0.91	1.8 с	465 мс	133 мс	31.1 с	9.7 с	1.6 с

Проведені дослідження показали, що розроблені методи гармонізації дозволяють досягати вищих показників гармонійності з використанням меншої кількості обчислень. Отримана модель (6) становить четвертий пункт наукової новизни.

Ключові слова: динамічний музичний синтез, генетичний метод, гармонізація, нейромережа, автокодувальник, генотипне представлення, область пошуку.

СПИСОК ПУБЛІКАЦІЙ ЗДОБУВАЧА

1. Komarov O., Galchonkov O., Nevrev A., Babilunga O. Consonant chord model of musical compositions for harmonizing melodies by a genetic algorithm. *Proceedings of Odessa Polytechnic University*, 3(56), 2018, с. 63 - 79, DOI 10.15276/opu.3.56.2018.07.
2. Komarov O., Galchonkov O., Nevrev A., Babilunga O. The use of a genetic algorithm for music harmonization of transitional fragments in computer games. *International Journal of Computing*, 19(3), 2020, с. 365-376. DOI: 10.47839/ijc.19.3.1886.
3. Komarov O. Reducing the search area of genetic algorithm using neural network autoencoder. *Herald of Advanced Information Technology*, 3(3), 2020, с. 113-124. DOI: 10.15276/hait03.2020.1
4. Komarov O., Galchonkov O., Boltenkov V. Reducing chromosome redundancy in genetic algorithms using autoencoder. *Proceedings of VII International scientific & technical conference «Informatics.Culture. Technology»*. Odessa, 2019, с. 77-78.
5. Комаров О., Болтъонков В. Система порівняння моделей представлення даних в тональній діатонічній музиці. *Proceedings of VII International scientific & technical conference «Informatics.Culture. Technology»*. Odessa, 2019. с. 75-76.
6. Komarov O., Wang Q. Improved genetic method of melodies harmonization with the help of peak values of error smoothing. *Матеріали Восьмої Міжнародної наукової конференції студентів та молодих вчених «Сучасні інформаційні технології 2018»*. Одеський національний політехнічний університет, Одеса, 2018, с. 120-121 .ISBN 978-617-7046-50-8
7. Komarov O., Boltenkov V. The Advanced Method for Melody Harmonization in Procedural Music Generation. *Proceedings of VI International scientific & technical conference «Informatics.Culture. Technology»*. Odessa, 2018.
8. Komarov O., Boltenkov V. Melody harmonization with form development in procedural music. *Proceedings of 3d International Conference on Computer Algebra and Information Technologies (CAIT-Odessa-2018)*. Одеський національний університет імені І.І. Мечникова, 2018, с. 164-166. ISBN 978-617-7613-18-2.

ABSTRACT

Komarov O.V. Models and methods of dynamic music synthesis based on a genetic approach. – Qualification scientific work in the form of the manuscript.

Thesis for the PhD degree in specialty 122 "Computer science". – Odessa Polytechnic State University, Ukraine, Odessa, 2021.

The **introduction** substantiates the urgency of solving the problem of developing methods and models of dynamic music synthesis with a high level of harmoniousness in limited computational performance. Dynamic music synthesis (DMS) is an automated process of forming musical compositions from pre-created melodic fragments due to their harmonization. It has been shown that DMS requires significant computing efficiency and usually requires consistency with external control processes, such as the actions of the user of the game application or the process of creating real-time media content. The object, subject, tasks and methods of research are defined; the connection with scientific programs and plans is shown; the scientific novelty and practical significance of the obtained results are given; the personal contribution of the applicant is covered.

In the **first section** of the dissertation is carried out the analysis of problems of development of dynamic musical synthesis systems with a high level of harmoniousness.

The essence of the DMS task is to create a composition by forming an accompaniment for a given melody. In this case, the composition is represented as a sequence of sets of simultaneously used sounds (chords), each of which includes one sound of the melody. Harmonization is performed on a set of melodies written by the composer in advance.

The existing methods of musical compositions synthesis, their advantages and disadvantages are analyzed. It is shown that the construction of dynamic music synthesis systems (DMSS) based on them is complicated because they synthesize music compositions that do not fully meet the requirements of harmony, and/or cannot meet these requirements in real time with limited computing performance. The

contradiction is formulated, which is solved in the dissertation - between the requirements for harmony and the possibility of providing them with limited time and computing performance.

An analysis of existing music systems for interactive multimedia projects. It is proposed to highlight those that are the most common and of particular interest: the diatonic music system and systems with regular modes (Fig. 1).

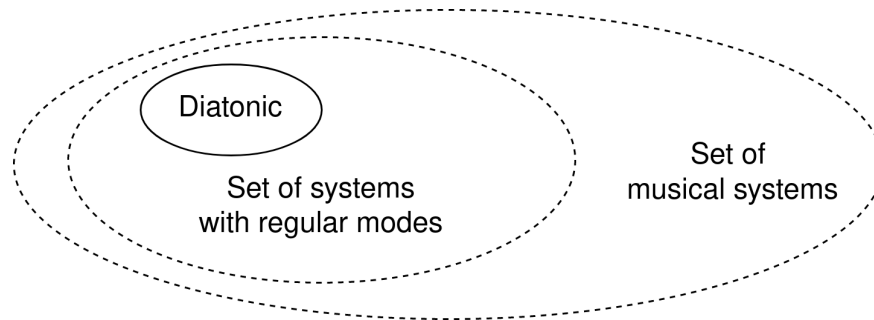


Figure 1 – Diagram of the relationship of common music systems.

It is shown that the most promising approach for the development of DMS methods is the evolutionary approach. Its application involves the development of genotypic representation models for the considered musical systems and the method of genetic harmonization.

Genetic methods for harmonization allow to form a set of equivalent, but different in content solutions with a high level of harmony, as the fitness function of the genetic method $H(c)$, that creates composition C , takes into account the concept of harmony, determined by a set of rules, and looks like:

$$H(C) = 1 - \frac{1}{Q(L-1)} \sum_{i=1}^{|F|} \sum_{j=1}^{L-1} q(f_i) f_i(\mathbf{A}_j, \mathbf{A}_{j+1}) \quad (1)$$

where L – number of chords (simultaneously used sounds) of composition, Q – sum of weights of all chord use rules, f_i – chord use rule, $i = 1, \dots, |F|$, $|F|$ – number of chord use rules, q – function of chord use weights, $\mathbf{A}_j$ – chord of composition, $j = 1, \dots, L-1$.

The disadvantage of existing genetic methods of harmonization is low efficiency due to the redundancy of the search area in which the search for a feasible solution is performed. The search area corresponds to the parameter space of genetic

representations, and its size depends on their coding. Usually, to ensure the search for the whole set of phenotypes, developers of genetic representations prefer completeness of representation and do not meet the requirement of legality. It is shown that increasing the legality of the genetic representation of a musical composition by taking into account the characteristics of a particular musical system for which harmonization is performed, can be a reserve of reducing the area and, consequently, the time of search.

The purpose of the dissertation research and its tasks are formulated.

In the *second section*, the consonant chord model of genotypic representation of a musical composition for the diatonic music system and the genetic method of harmonization with its use were developed. When developing a model to increase the legality of genotypic representation, it is proposed to apply the principle of feasible completeness, which assumes the existence of genotypes only for feasible solutions by increasing the structural generality of the representation. Structural generality is determined by the level of abstraction of musical concepts of the representation parameters. There levels of structural generality are shown: sound, mode and chord.

Within the framework of the highest, chord level of structural generality on the basis of the research of structure of diatonic system's triad chord, a set of its parameters is proposed:

- function $f \in \{1, 2, 3\}$ defines the beginning of tertian order, its degree composition and role in chord progression;
- inversion $i \in \{1, 2, 3\}$ defines the beginning of voice order and the degree of lowest sound in the chord;
- placement $p \in \{1, 2\}$ determines the relative position of middle voices;
- repeat $r \in \{1, 3\}$ determines the repeated degree.

Based on this parameter set, a consonant chord model of genotypic representation of $\mathbf{C}^G$ musical compositions in the form of progression of chord images $\mathbf{A}_j^G = (m_j, f_j, i_j, p_j, r_j)$ is proposed:

$$\mathbf{C}^G = \begin{bmatrix} m_1 & m_2 & \dots & m_L \\ f_1 & f_2 & \dots & f_L \\ i_1 & i_2 & \dots & i_L \\ p_1 & p_2 & \dots & p_L \\ r_1 & r_2 & \dots & r_L \end{bmatrix} \quad (2)$$

where m_j – melody sounds, $j = 1, \dots, L$, L – melody length.

The phenotypic representation of the musical composition $\mathbf{C}$ for the genetic harmonization method with fitness function (1) is a matrix that is a sequence of L of chords $\mathbf{A}_j$:

$$\mathbf{C} = \{\mathbf{A}_j : \mathbf{A}_j = \{P_{kj} | k = 1, \dots, 4\}\} \quad (3)$$

where P_{kj} – pitch values of individual sounds.

For decoding of the model (2) to phenotype (3) is proposed function *pitch*, which map each j -th column of $\mathbf{C}^G$ to chord $\mathbf{A}_j$:

$$\mathbf{A}_j = \text{pitch}(\text{order}(\text{degrees}(m_j, f_j), i_j, p_j, r_j), m_j) \quad (4)$$

where *degrees* – a function that forms a vector of chord degrees, arranged in tertian order, *order* – a function that forms a vector of chord degrees corresponding to the chord sounds, arranged in a vocal order.

$$\text{degrees}(m, f) = (\delta^d : \delta^d = (7 + \delta(m) - 2(f - d)) \bmod 7)$$

where δ^d – mode degree, $d \in \{1, 2, 3\}$, δ – function of correspondence between pitches and degrees of mode, *mod* – modulo operation.

Function *order*(Δ, i, p, r) arranges and expands created degree vector of tertian order Δ , forming new degree vector in voice order $\Delta_p = (\delta_1, \delta_2, \delta_3, \delta_4)$:

$f \neq i$	$f = i = 2$	$f = i \neq 2$
$\delta_1 = \delta^f$ $\delta_{1+p} = \Delta \setminus \{\delta^f, \delta^i\}$ $\delta_{4-p} = \delta^r$ $\delta_4 = \delta^i$	$\delta_1 = \delta^f$ $\delta_2 = \delta^p$ $\delta_3 = \delta^{3-p}$ $\delta_4 = \delta^r$	$\delta_1 = \delta^f$ $\delta_{1+p} = \Delta \setminus \{\delta^f, \delta^2\}$ $\delta_{4-p} = \delta^2$ $\delta_4 = \delta^i$

pitch function transforms this degree vector Δ_P into pitch vector $\mathbf{A}_j$ based on the melody sound pitch m :

$$pitch(\Delta_P, m) = \left\{ P_k : P_k = \begin{cases} m & k = 1 \\ \delta^{-1}(\delta_k, \omega(P_{k-1})) & P_k < P_{k-1} \\ \delta^{-1}(\delta_k, \omega(P_{k-1}) - 1) & P_k \geq P_{k-1} \end{cases} \right.$$

where δ^{-1} – function of correspondence between the mode degree δ_k and octave ω_k of sounds and its pitch, $k \in \{1, 2, 3, 4\}$.

Thus, the first point of scientific novelty was obtained: for the first time a consonant chord model of genotypic representation of musical compositions (2) was proposed, the feature of which is compliance with the principle of feasible completeness due to the use of chord level of structural generality, which allowed to reduce the search space in genetic method of harmonization.

Based on model (2), a genetic method of harmonization is proposed, which involves the following steps:

1. Form the initial population of N individuals in accordance with the selected model of representation, for example, for the diatonic system – $\mathbf{C}^G$ (2).

2. Set the iteration counter to 0.

3. To calculate the fitness function $H(C)$ (1) decode the population into a phenotypic representation (3) with correspondence to the model of the selected music system. For example, in the case of diatonic – using function (4).

4. If the value of fitness function $H(C)$ (1) is not less than a predetermined threshold or the value of the iteration counter is greater than a predetermined threshold, finish. Otherwise, increase the iteration counter by 1 and go to the next step.

5. Select n most harmonious individuals of the population.

6. Create a new population of size N by a single-point crossover between each pair of n selected individuals.

7. Mutate the population, with a chance of p for each individual.

8. Return to step 2.

Research has shown that the use of the proposed genetic method of harmonization can reduce the number of iterations required to achieve a level of harmony of 0.9 by 4 times compared with the method without using model (2).

The proposed method can be used for harmonization with other genotypic representation models and is the second point of scientific novelty.

In the *third section* the method for creating genotypic representation models of musical compositions for an arbitrary music system and a system with a regular mode based on the use of a neural network autoencoder for decoding chord parameters is developed.

A neural network autoencoder is a multilayer neural network of symmetrical architecture with a narrow central layer that learns to form on the central layer a lower-dimensional image of input data.

A model of genotypic representation $\mathbf{C}^E$ of musical composition for an arbitrary music system and a system with a regular mode is proposed in the form of a sequence of chords images $\mathbf{A}_j^E$:

$$\mathbf{C}^E = \{\mathbf{A}_j^E : \mathbf{A}_j^E = \{m_j\} \cup \{a_{ej} : a_{ej} \in \{0, 1\}\}\} \quad (5)$$

where m_j – pitch of melody sound, $j = 1, \dots, L$, L – melody length, a_{ej} – chord parameters, $e = 1, \dots, E$, E – the selected encoding size of the chord.

Based on the formed sample of chords of a given music system, an autoencoder with a central layer of size E is trained. During the initialization of the genetic harmonization method, a population of individuals of representation (5) is created, where parameters a_{ej} are filled with random binary values. At the stage of individual evaluation, the created population is decoded into a phenotypic representation (3).

For an arbitrary music system has been developed the method of creating genotypic representation models of musical compositions by training an autoencoder, which involves the following steps:

1. Transform each vector of pitches of the chord $\mathbf{A}_s^\sigma = \{P_{ks}\}$ from the sample to the vector of intervals $\mathbf{A}_s^A$:

$$\mathbf{A}_s^\Delta = \{\Delta P_{ks} : \Delta P_{ks} = P_{(k-1)s} - P_{ks}\}$$

where $k = 2, \dots, 4, s = 1, \dots, S, S$ – sample size.

2. Build a histogram of ΔP_{ks} values. Based on a predetermined allowable frequency of interval use α , find the smallest interval ΔP_m for which $p(\Delta P_m) < \alpha$.

3. Remove from the sample vectors $\mathbf{A}_s^\Delta$, which contain $\Delta P_{ks} \geq \Delta P_m$.

4. Transform the values of the intervals to binary vectors of length b , where b is a natural number corresponding to the following inequality:

$$2^{b-1} < \Delta P_m \leq 2^b$$

5. Concatenate binary vectors of intervals of one chord.

6. Train the autoencoder with the central layer of the size E with use of the received set of vectors.

7. Form a decoder by removing the input and subsequent layers up to the central from the autoencoder and replacing the central layer with the input layer.

During the fitness function calculation, the binary part of each chord of the form (5) is processed by the decoder, divided into separate vectors, which are debinarized to decimal values ΔP_{ks} and converted to the pitch values P_{ks} , starting from the value P_{la} , forming a phenotypic representation of each chord.

The model $\mathbf{C}^E$ can be used in genetic harmonization methods. It has been experimentally proven that the use of the proposed model with such decoding allows to reduce the number of iterations required to achieve a harmony level of 0.9 for an arbitrary music system by 1.8 times compared to the method without the model $\mathbf{C}^E$.

The method of creating genotypic representation models of compositions for music systems with regular mode differs from the previous one by calculating the degree intervals instead of pitch intervals. To do this, the reconstruction of the system mode is performed by following steps:

1. Build a histogram of the distribution of the pitches P_{ks} used in the chords $\mathbf{A}_s^\sigma$ by the pitch classes, where pitch class is the equivalence class to which belong any two pitches P_φ i P_ϕ , if the following is valid:

$$P_\varphi = 2^n P_\phi, n \in \mathbb{Z}$$

2. Based on the accepted allowable frequency of pitch class use γ , remove pitch classes P^{Ci} , for which $p(P^{Ci}) < \gamma$.

3. Set the order on the pitch classes by defining the function:

$$\delta : P \rightarrow D, D \subset \mathbb{N}$$

where P – set of all pitches, D – set of all mode degrees.

Degree intervals can be calculated as follows:

$$P_{1s} \ominus P_{2s} = (\delta(P_{1s}) - \delta(P_{2s})) \bmod |D|$$

Steps 2 and 3 of the method for an arbitrary music system can be omitted, since b corresponds to the following inequality:

$$2^{b-1} < |D| \leq 2^b$$

Because the value of b in this method is less than in the method for an arbitrary music system, the size of the encoding E is reduced, reducing the parameter space of the chord.

During the fitness function calculation, the binary part of each chord of the form (5) is processed by a decoder, divided into separate vectors, which are debinarized to decimal values, and then converted into values of chord degrees arranged in a voice order, forming a set Δ_P . The pitch vector of the phenotype is formed by function (4), calculated on Δ_P and P_{Is} .

It is experimentally proved that the use of the proposed model with such decoding in the developed harmonization method allows to reduce the number of iterations required to achieve a harmony level of 0.9 in systems with regular mode by 3.5 times compared to the method without model C^E .

Thus, the third point of scientific novelty is obtained: for the first time a method of creating models of genotypic representation of musical compositions for arbitrary music system and system with regular mode, the feature of which is compliance with the principle of feasible completeness due to the use of neural autoencoder for achieving structural generality, which allowed to reduce the search space in genetic method of harmonization.

In the *fourth section* the structure of the dynamic musical synthesis system is developed (fig. 2), the model of an estimation of necessary computing performance of DMSS on the basis of the genetic approach is proposed.

The presented system connects the states of the system in a form of automata, which shows the possible transitions from one state to another. The external process, which is consistent with the work of DMSS, acts as an input signal for this automata. The repertoire contains musical themes that are correlated with the states of the automata. The transition area generator generates a melody for harmonization by combining a fragment of the current theme with the next fragment of the current and initial fragments of themes of adjacent states. Depending on the music system in which the melody is formed, the genetic method configurator selects the rules of harmony and the corresponding models of representation - a consonant chord model (2) or any number of models (5) created using an autoencoder. The genetic method harmonizes the compositions for all possible scenarios of transitions for the current state, adding them to the appropriate bank in the form of phenotypes (3).

When a signal with information about the current state is received, the system plays one of the pre-harmonized scenarios, which corresponds to the current state, and begins to harmonize the scenarios after updating the state of the automata.

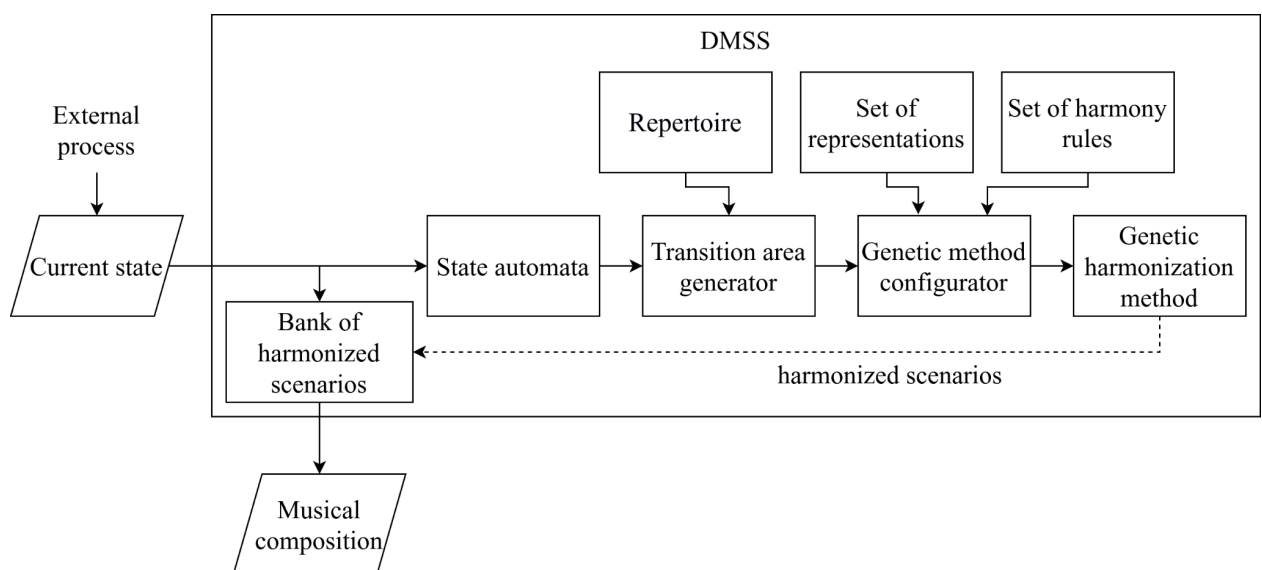


Figure 2 – Structural and functional scheme of DMSS.

Based on the relationships between the maximum allowable duration of harmonization, the number of harmonized chords, and the bars of melodies, two different modes of DMSS operation are proposed: high-speed and high-quality.

A model for calculating the computational requirements of the developed DMSS is proposed. The required computing power M to harmonize the repertoire of scenarios in t_{min} seconds is:

$$M = \sum_{\omega}^{\Omega} \frac{IN(L_{\omega} - 1)\lambda_{\omega}}{t_{min}} \quad (6)$$

where Ω – the largest set of transition areas for the current state of the system, ω – the transition area of a particular scenario, I – number of iterations of the genetic algorithm, N – population size, L_{ω} – the melody length of the transition area ω , λ_{ω} – the number of operations required to calculate the fitness function for one pair of chords on the model associated with the transition area ω .

Model (6) allowed to analytically determine the required computing power for DMSS based on the genetic method in given time constraints using the developed genotypic representations of the chord level of structural generality (2) and (5) and the model of representation of the lower - mode - level (table 1 – table 3).

Table 1 – Harmony characteristics and computing power required to harmonize one scenario in high-speed mode.

H	Mode level model		Model C^E (5) for regular mode system		Model C^G (2)	
	I	$M^I, 10^6$ FLOPS	I	$M^I, 10^6$ FLOPS	I	$M^I, 10^6$ FLOPS
0.61	1	0.208	1	0.065	1	0.049
0.7	3	0.697	1	0.065	1	0.049
0.82	6	1.659	3	0.225	1	0.049
0.91	13	3.647	5	0.388	4	0.212
0.94	31	8.972	8	0.635	5	0.269

Table 2 – Harmony characteristics and computing power required to harmonize one scenario in high-quality mode.

H	Mode level model		Model $C^E(5)$ for regular mode system		Model $C^G(2)$	
	I	$M^I, 10^6$ FLOPS	I	$M^I, 10^6$ FLOPS	I	$M^I, 10^6$ FLOPS
0.61	3	2.999	1	0.402	1	0.13
0.7	6	7.108	3	1.381	1	0.13
0.82	9	11.772	6	3.156	1	0.13
0.91	14	19.787	9	4.892	6	0.874
0.97	35	52.435	28	15.959	20	3.045

The comparison of harmonization durations was performed on a hardware platform with characteristics: Intel(R) Xeon(R) CPU @ 2.20GHz (1 core, 2 threads, 32 bit, cache 512 kb), 13 GB RAM. The test was performed in two modes for two levels of harmony.

Table 3 – Time required to harmonize one scenario in different modes.

H	High-speed			High-quality		
	Mode level model	Model $C^E(5)$ for regular mode	Model $C^G(2)$	Mode level model	Model $C^E(5)$ for regular mode	Model $C^G(2)$
0.7	414 mc	93 mc	34 mc	13.6 c	3.9 c	269 mc
0.91	1.8 c	465 mc	133 mc	31.1 c	9.7 c	1.6 c

Research have shown that the developed methods of harmonization allow to achieve higher harmony with the use of fewer calculations. The obtained model (6) is the fourth point of scientific novelty.

Keywords: *dynamic music synthesis, genetic method, harmonization, neural network, autoencoder, genotypic representation, search area.*