YEGAH MUSICOLOGY JOURNAL
HTTP://DERGIPARK.ORG.TR/PUB/YEGAH
e-ISSN: 2792-0178

Makalenin Türü / Article Type	: Araştırma Makalesi/ Research Article
Geliş Tarihi / Date Received	: 20.08.2024
Kabul Tarihi / Date Accepted	: 09.09.2024
Yayın Tarihi / Date Published	: 23.09.2024
DOI	: https://doi.org/10.51576/ymd.1536267
e-ISSN	: 2792-0178

İntihal/Plagiarism: Bu makale, en az iki hakem tarafından incelenmiş ve intihal içermediği teyit edilmiştir. / This article has been reviewed by at least two referees and confirmed to include no plagiarism.

RESEARCH ON THE CONSTRUCTION OF AI COMPOSITION SYSTEM BASED ON HMMS

YANG, Weijia[1]
LEE, Inho[2]

ABSTRACT

The era of Artificial Intelligence (AI) in music has truly arrived. Since the groundbreaking success of "Genesis," the first AI-composed album released in 2016 by the world's inaugural Artificial Intelligence Virtual Artist (Aiva), the music industry has witnessed an AI revolution. This momentum continued with the remarkable launch of Suno AI V3 in early 2024, capable of generating publishable music within minutes based on simple prompts. AI composition has now become a focal point of interest for scholars around the globe. Currently, there are many automatic composition algorithms in the interdisciplinary research field of musicology and computer science, and Markov chain is a representative automatic accompaniment algorithm for intelligent computer systems. With the advent of the big data era, Markov chain, an automatic composition method based on probability theory, has gradually been ignored by everyone, but research on artificial intelligence music generation based on Hidden Markov model (HMM) is still very valuable. This

[1] Dr. Weijia Yang, Department of Global Music, Graduate School of Kyonggi University, 24, Kyonggidaero 9-gil, Seodaemun-gu 03746, Republic of Korea, vjkisa520@gmail.com, https://orcid.org/0009-0007-8400-9449 .

[2] Corresponding Author, Assist.Prof.Dr. Inho LEE, Department of Applied Music, Kyonggi University, 24, Kyonggidaero 9-gil, Seodaemun-gu 03746, Republic of Korea, soulinno@kgu.ac.kr, https://orcid.org/0000-0001-7096-0670 .

paper proposes a method for constructing an AI composition system based on the HMM. This system achieves the goal of automatically generating accompaniment music from score data. The proposed system has achieved relatively stable results in the generation of musical elements such as form and harmony, accompaniment texture, and instrumentation, and has scored well in evaluation experiments. This study hopes to explore the technical boundaries of musicology research in the AI era through the construction of algorithmic composition engineering cases in an empirical way.

Keywords: Automatic accompaniment algorithm, Artificial Intelligence, composition, popular music, Hidden Markov Model.

HMMS TEMELLİ YAPAY ZEKÂ BESTECİLİĞİ SİSTEMİ İNŞASI ÜZERİNE ARAŞTIRMA

ÖZ

Müzikte Yapay Zekâ (YZ) çağı gerçekten başladı. Dünyanın ilk YZ Sanal Sanatçısı (Aiva) tarafından 2016 yılında yayımlanan ve büyük bir başarı elde eden “Genesis” adlı albümden, 2024’ün başlarında yalnızca birkaç kelimelik bir komutla dakikalar içinde yayımlanmaya hazır müzik üretebilen Suno AI V3’ün dikkat çekici çıkışına kadar, müzik endüstrisi bir YZ devrimi yaşamaktadır. YZ ile beste yapımı, artık dünyanın dört bir yanındaki akademisyenler için ilgi odağı haline gelmiştir. Şu anda, müzikoloji ve bilgisayar bilimi disiplinler arası araştırma alanında birçok otomatik besteleme algoritması bulunmaktadır ve Markov zinciri, zeki bilgisayar sistemleri için temsil edici bir otomatik eşlik algoritmasıdır. Büyük veri çağının gelmesiyle birlikte, olasılık teorisine dayalı bir otomatik besteleme yöntemi olan Markov zinciri yavaş yavaş göz ardı edilmeye başlanmıştır. Ancak, Gizli Markov Modeli (HMM) temelli YZ müzik üretimi üzerine yapılan araştırmalar hala çok değerlidir. Bu makale, popüler müzik tarzı için HMM’ye dayalı bir YZ beste sistemi inşa etme yöntemini önermektedir. Bu sistem, nota verilerinden otomatik olarak eşlik müziği üretme hedefini gerçekleştirmektedir. Önerilen sistem, form ve uyum, eşlik dokusu ve enstrümantasyon gibi müzikal unsurların üretilmesinde görece istikrarlı sonuçlar elde etmiş ve değerlendirme deneylerinde iyi notlar almıştır. Bu çalışma, YZ çağında algoritmik besteleme mühendisliği örneklerinin inşası yoluyla müzikoloji araştırmalarının teknik sınırlarını deneysel bir şekilde keşfetmeyi amaçlamaktadır.

Anahtar kelimeler: Otomatik eşlik algoritması, Yapay Zekâ, beste, popüler müzik, Hidden Markov Model.

INTRODUCTION

Artificial Intelligence (AI) has emerged as a new scientific research field in recent years within the realm of computer science. The research on (Minsky, 2007) and (Gurkaynak et al., 2016) indicates that it encompasses the relevant theories, methods, technologies, and applications for constructing human intelligence systems and has consistently been a focal point of research. Professor (Zhu, 2017) and professor (Tang, 2021) express that the development of artificial intelligence is built upon the mutual permeation of foundational disciplines such as computer science, philosophy, and linguistics. (Farbood & Schöner, 2001), (Allan & Williams, 2004) and (Bell, 2011), among others, multiple studies have pointed out that AI composition stands as a representative intersection between computer science and musicology. (Sako et al., 2014)'s research points out that the markov chain has been a foundational aspect of AI composition research. As early as (Simon et al., 2008)'s research, it particularly emphasized the research perspective of "following limited melodic score data and generating accompaniment in real-time". Simultaneously, both studies on (Luque, 2009) and (Fernández & Vico, 2013) indicate that AI composition can be regarded as an advanced method of intelligent accompaniment and an important part of the algorithmic composition system. The construction of algorithms based on music composition theory is the core of its technology.

The artificial intelligence automatic accompaniment has developed very rapidly in recent years due to the progress of science and technology (Oliwa, 2008), especially on the international level, where artificial intelligence accompaniment already has a relatively systematic approach (Xia & Dannenberg, 2015). The following research case on AI Composition and Hidden Markov Model (HMM) is the basic reference for this study. This study also focuses on algorithm modeling research on this topic.

About the AI Composition, (Dannenberg & Grubb, 1994) published intelligent accompaniment of real-time music, marking an early initiation of real-time intelligent accompaniment experiments. (Bäckman, 2009)'s optimal harmony progression and (Doush & Sawalha, 2020)'s accompaniment system based on polyphonic rock music are both anchored in the performer's 'score,' analyzed by a computer program, and then synthesized into the accompaniment. Examples such as (Thom,

2000) and (Linson et al., 2012)'s improvisation system using artificial intelligence are also relatively well-known.

About the HMM, Markov Chain and Hidden Markov Chain are crucial algorithms in the AI composition domain, especially in the field of intelligent accompaniment, where these algorithms are frequently employed for data analysis or function matching. The rule-based knowledge base system algorithm, also known as an expert system (Dannenberg & Hu, 2003), such as tonal harmony rules (Chong & Ding, 2014), can be cumbersome due to a large number of rules and a huge database. (Rohrmeier & Graepel, 2012) and (Herremans et al., 2017) built a database based on Bach Chorales and performed harmony matching operations based on the Hidden Markov Chain. Additionally, (Shan & Chiu, 2010) and (Zarro & Anwer, 2017), based on the alignment work of Palestrina, employed the Markov Chain for harmony matching operations. Furthermore, the accompaniment system of (Orio, 2001) and (Jordanous & Smaill, 2008) also utilizes Hidden Markov Models.

In the aforementioned research, the Hidden Markov algorithm is employed to predict the next harmony or texture. Whether the music scores follow the algorithmic composition system or popular tools like "(Band-in-a-Box, 2024)" and "(Music Memos, 2024)", they all represent the standardization trend of the composition process. The methods used in the mentioned literature mostly rely on the theory of artificial intelligence model. After analyzing the melody of the song, the result of the generated harmony and accompaniment is executed. Therefore, this article aims to implement the AI Composition System based on HMM (HMM-AICS) in a more innovative and convenient way. This experimental work aligns with the trend by focusing on "AI Composition" and "HMM."

This article proposes HMM-AICS based on the existing research on AI composition, combined with the music theory of pop music style and the application of orchestration. The first step involves melody identification, analyzing the tonality, phrases, rhythm, tempo, and other features of the melody. The second and third steps focus on program theme content, and the final part involves accompaniment output. The system is divided into four parts: music element analysis, harmony matching, orchestration matching, and accompaniment output, as illustrated in Figure 1. Based on this structure, this study completed the construction of a database based on music theory, and further integrated the HMM into this system. In the last part of this study, the experimental verification of musical works was also set up.

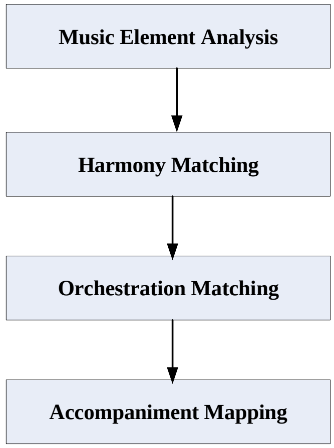


Figure 1. The Structure of HMM-AICS.

METHOD

Popular Music Style Database Construction

The database of the proposed HMM-AICS is based on music theory, which can make the accompaniment more consistent with the rules of music creation and optimize the operation of the program. There are two concepts in pop music, namely broad and narrow. In this study, it specifically refers to the “non-classical” commercial music style. In the accompaniment of pop music, notes, harmony, and rhythm are crucial structural components (Wright et al., 2012). This section revolves around the construction of three databases: “Notes, Harmony, and Rhythm.” The Notes Basic Database explains fundamental data about the music, such as tone color, pitch, and duration. The Pitch Database specifically focuses on major and minor tonalities of harmony or melody, as well as other special pitches. The Harmony Database serves as the foundational corpus for automatic accompaniment, including common triads, seventh chords, suspended chords, etc. Additionally, it can enhance the richness of harmonic vocabulary through algorithms related to random probabilities. The Rhythm Database serves as a reference for the musical bass and drum sections. Based on note duration and values, musical rhythms are composed, and after grouping several rhythms, they are stored in the rhythm database.

Notes database

Note data utilizes universal MIDI information as the architectural standard. This paper defines the representation method for each note, including P (pitch), D (duration), R (rhythm), V (velocity), T (timbre), as shown in Equations (1) to (6). Here, "a" represents the MIDI pitch number, "b" represents the MIDI note duration, "c" represents the combined MIDI rhythm grouping of "a" and "b", "d" represents the MIDI velocity, and "e" represents the MIDI instrument number.

$$\text{Tone} = F \quad (P, V, D, T) \tag{1}$$

$$\text{P(Pitch)} = 12k + a \tag{2}$$

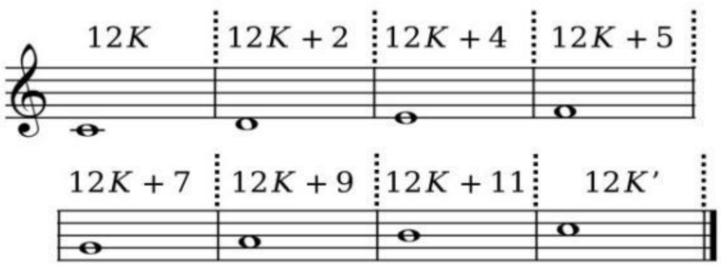
$$\text{D(Duration)} = b \tag{3}$$

$$\text{R(Rhythm)} = c \tag{4}$$

$$\text{V(Velocity)} = d \tag{5}$$

$$\text{T(Timbre)} = e \tag{6}$$

Pitch is the musical element that needs to be defined first. By defining pitch, the arrangement of scales can be determined. Based on the different arrangements of scales, different modes and tonalities can be formed. Scales and modes are the foundation of harmony. The correspondence between pitch and scales is shown in Score 1. Here, "k" represents the octave pitch position, a represents the relative pitch position within the same octave, and "$12k + a$" can accurately represent any pitch position. It is important to note that both k and k' are integers and are adjacent octaves.



Score 1. Correspondence chart of natural major scale and pitch.

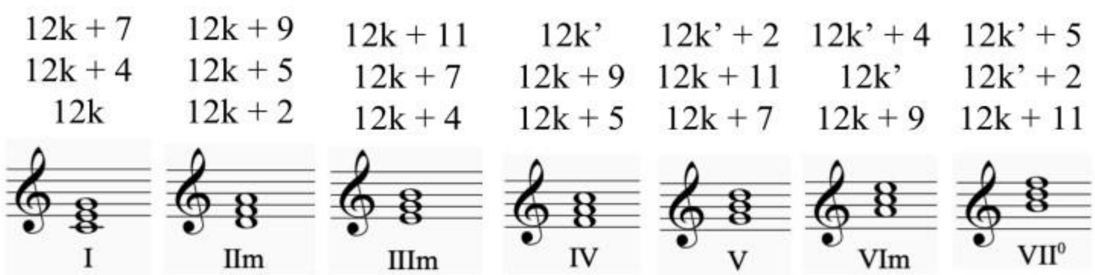
Six major tonality databases, including natural, melodic, and harmonic scales, can all be defined and stored using this method. Other commonly used scales such as blues, medieval, and ethnic scales can also be extended based on this approach. By considering pitch and scales, chords can be deduced, completing the establishment of the harmonic progression.

Harmony database

Setting harmonic data involves calculating adaptable harmonic combinations based on the identified melody. This section focuses on summarizing the mapping methods between harmonic theory and data. Due to program limitations, this paper uses the common tonality in popular music, specifically the harmonic major tonality, as a case study. Starting from functional harmony, Table 1 reflects the basic harmonic functional attributes. From popular music to classical music, harmonic movement unfolds as T (Tonic) - S (Subdominant) - D (Dominant). Therefore, the harmonic database is extended based on these three fundamental functional groups. Comparing with Score 2, each chord corresponds to pitch expression parameters based on the natural major tonality. The subsequent variations in harmonic color and progressions are derived from this.

Tonic Function	I	III_m
Subdominant Function	IV	II_m
Dominant Function	V	III_m

Table 1. Basic Harmony Function Property Sheet.



Score 2. The Corresponding Position is the Chord Storage Parameter Based on Natural Major.

(1) Chord Color Transformation

Expand harmonic color beyond basic functional harmony. Functional harmony only satisfies basic harmonic combinations, and by adding or replacing pitches, harmonic color can be further enriched. As shown in Table 2, all colors can be classified into major chords, minor chords, suspended chords, and dominant seventh chords, these being the four basic colors. Then, by adding notes, 13 commonly used compound chords such as seventh chords and ninth chords can be formed. Eight alternative chords can also be obtained by replacing the third note.

Basic Color	Augmented Chord	Altered Chord (3rd Tone)
Major	6th, M7, M9, add2	sus2, sus4
Minor	6th, m7, m9, m11, add2	m, sus4
Suspended	sus2	None
Dominant 7	9th, 13th, #9th	6th, 7th sus4, 11th

Table 2. The Basic Chord Color and Conversion Table.

Map the chords of the four basic colors. By replacing and adding notes, a rich harmonic color palette can be achieved, and the data can be represented by three types of actions: replacing pitches, adding pitches, and adding materials for compound chords. It can be understood as changing the third note based on the triad, or directly using seventh chords, ninth chords, and eleventh chords. Refer to the three tables below, which are further extensions of Table 2. The left column represents the indication of replacing, adding, or stacking chords, while the right column expresses pitch or chord codes. In this context, "*a*" represents the third note, belonging to a pitch position within an octave, and "*b*" is an additional pitch, excluding the pitch of the triad, in the same octave range.

Alter the 3rd Tone / 12k + a	a = {2, 3, 4, 5}
Add a Pitch / 12k + b	b = {2, 8, 11}
9th Chord	12k + 11; 12k + 13

Table 3. Major Chord Basic Harmony Color Deformation Representative Code.

Alter the 3rd Tone / 12k + a	a = {3, 4, 5}
Add a Pitch / 12k + b	b = {2, 8, 11}
9th Chord	12k + 11; 12k + 13
11th Chord	12k + 11; 12k + 13; 12k + 17

Table 4. Minor Chord Basic Harmony Color Deformation Representative Code.

Suspended Chord: Alter the 3rd Tone / 12k + a	a = {2}
Dominant 7 Chord: Alter the 3rd Tone / 12k + a	a = {9, 5}
9th Chord / 12k + b; 12k + c	b = {11}; c = {2, 3}
13th Chord	12k + 11; 12k + 2; 12k + 5; 12k + 9

Table 5. Suspended Chord and Dominant Chord Basic Harmony Color Deformation Representative Code.

(2) Harmony Progression Settings

Based on the harmony degrees, summarize the matrix related to harmony degrees. The harmony degree matrix represents the probability of each harmony degree appearing in the song. This probability is not fixed and is influenced by factors such as the overall musical structure, rhythm, and cadence. Throughout the process of harmony progression, in addition to considering the degree of melodic matching, additional explanations for the overall direction of harmony are needed. The probability of harmony progression in major keys is shown in Table 6. The matrix at the intersection of the X and Y axes is organized based on the frequency of occurrence of various degree harmonies in popular music, and similar methods can be used to define probability matrices for other modes. Based on the harmony database and chord probabilities, matching harmonies can be calculated. After determining the harmony degrees, the harmonic colors can be expanded, but it is important to adhere to common harmonic patterns in popular music. The probability of chromatic harmony substitution can fluctuate around 30%, and it is important to limit the number of chromatic harmony substitutions to three times in the same musical section.

	VII°	**VIm**	**V**	**IV**	**IIIm**	**IIm**	**I**
I	0.3	0	0.3	0	0	0.3	0
IIm	0.2	0.3	0.06	0.2	0.2	0.02	0.01
IIIm	0.3	0.2	0.2	0.1	0.05	0.1	0.05
IV	0.3	0.15	0.15	0.08	0.3	0.01	0.01
V	0.02	0.2	0.06	0.2	0.2	0.3	0.02
VIm	0.2	0.06	0.2	0.2	0.3	0.03	0.01
VII°	0.07	0.18	0.18	0.18	0.18	0.18	0.03

Table 6. The Major Key Chord Progression Probability Transition Matrix based on Popular Music Theory.

Rhythm database

Rhythmic components determine the rhythm of a pop song. Based on the formulas 1 and 3 mentioned earlier, musical rhythms can be compiled according to the defined durations, and several rhythms can be grouped and stored in the rhythm database. Duration is the primary focus. The duration of notes is divided based on time, and rhythm and beats are based on this division. Duration is measured in milliseconds. In regular calculations, a rhythm of 60 BPM (beats per minute) means 60 vibrations per 1000 milliseconds (1 minute), and other speeds can be calculated accordingly. If a whole note corresponds to 4000 milliseconds, the note values in Table 7 can be calculated.

Note	Whole	Half	Quarter	8th	16th	32th
Duration (ms)	4000	2000	1000	500	250	125

Table 7. The Numerical List of the Note Duration.

Based on the duration, the rhythmic section for the bass can be composed. The bass is the fundamental part of harmony, defining not only the harmonic functions of the music (root position, inversion) but also establishing the foundation of the texture. The bass's range is preferably between 30 and 50 (MIDI pitch range). The system classifies the bass into three types: regular bass, melodic bass, and patterned bass. Regular bass uses the root (sometimes the fifth) of the chord as the bass. Melodic bass is often created by adding passing tones, introducing a melodic quality to the bass line. Patterned bass is based on the decomposition and semi-decomposition structure of the chord, also relying on the root and fifth. Each type of bass is associated with 10 common rhythmic patterns, serving as templates for the system. In units of short phrases, these rhythmic patterns are called upon, looped within the same section, and undergo slight variations.

The bass voice database can be synchronized with the percussion voice database. When consolidating the databases, the drum groups in the percussion voice database are sorted by rhythmic complexity. Based on the bass voice, similar bass groups are selected. The bass voice database includes multiple categories of rhythmic combinations. Once the pitch is determined, rhythmic combinations are automatically matched and stored in a temporary buffer, awaiting subsequent calls from the system.

In summary, this section focuses on the music theory database, revolving around the construction of the "notes, harmony, and rhythm" databases. It specifically defines algorithms for musical elements such as pitch, rhythm, and duration. Regarding dynamics and timbre, MIDI CC data can be temporarily utilized, and future extensions for sound quality can be achieved through third-party VST sources. This study aims to discuss musical content, avoiding further exploration of sound quality. The database is not only applicable to the analysis of musical elements but can also be invoked in harmony matching, orchestrator matching, and other areas.

RESULTS

AI Composition System for Popular Music

The proposed HMM-AICS architecture includes the harmony matching part and polyphonic matching part. The harmony matching part is composed of fundamental harmony matching, altered chord matching, and inversion/origin. The polyphonic matching part is composed of percussion part matching, low-pitched part matching, harmony texture part matching, and secondary melody part matching. As shown in Figure 2, the proposed AISFSPM can automatically generate chords and accompaniment textures based on MIDI file data. The harmony arrangement process is mainly shown in Figure 3, which is also the accompaniment generation database application process.

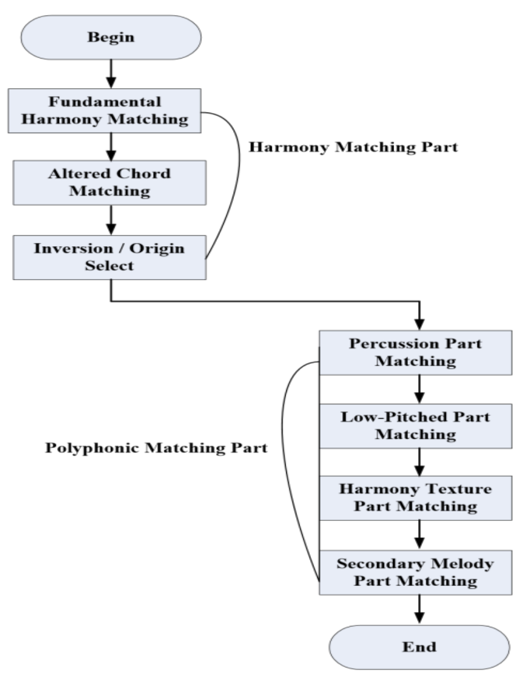


Figure 2. The Proposed HMM-AICS Architecture.

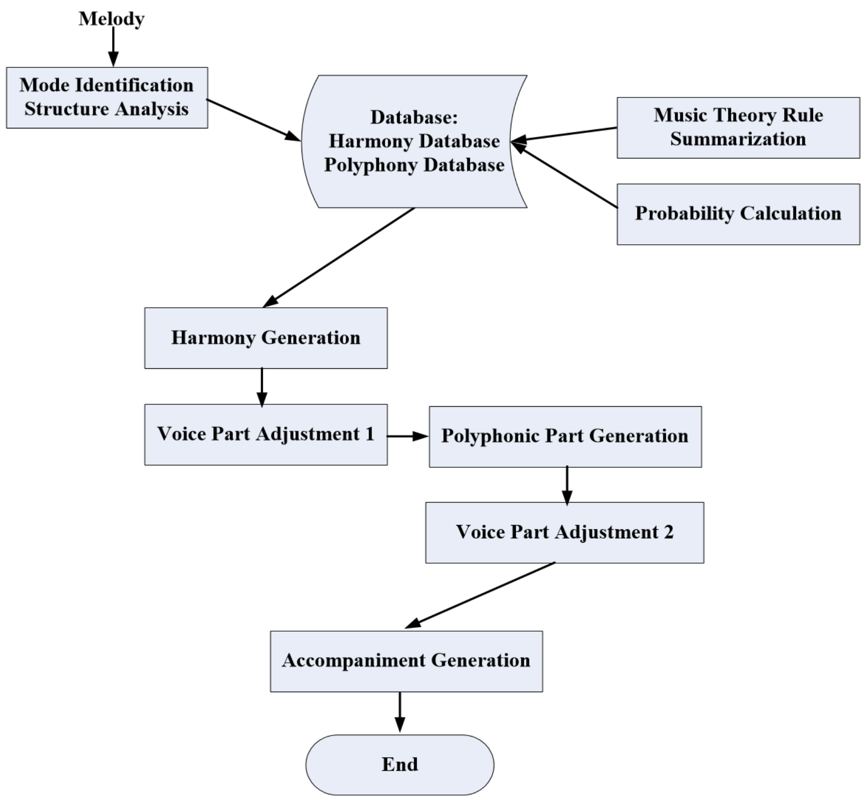


Figure 3. The Flow Chart of Accompaniment Generation Database Application.

Hidden Markov Model Harmony Matching Analysis

Hidden Markov Model (HMM) includes 2 state classes and 3 probability matrixes. The probability matrixes includes hidden state S, observable state O, and initial state probability matrix π, hidden state transition matrix A, observed state transition probability matrix B. HMM can be expressed as λ=(A, B, π) three elements, or λ=(N, M, A, B, π) five elements.

In harmony matching, each harmony states Dn (D1, D2, D3, D4,…Dn) can be matched from the melody, and the HMM can calculate the visible final state. The five elements corresponding to harmony matching are a hidden state within the same measure available chorus state. The initial state matrix is the hidden state matrix, and is retrieved from the database, can finally obtain the matched chord output.

(1) Hidden State Class N

S is the harmony state class, which can be directly called from the database. The n states are $\theta 1, \theta 2, \dots \theta n$, and at time t , the Markov state is qt, and $qt \in \{\theta 1, \theta 2, \dots \theta n\}$.

(2) Observable State M

The melody trend can be analyzed from the database, and divided m into $t + 1$ measures, where t duration determines the chord density, change a chord every t. This time is the observable state at time t, and the number of possible chords corresponding to each state. The number of chords matches the corresponding state in the database. Let n states be $V1, V2, \ldots Vn$. The observable state at time t is $Ot\ \{V1, V2, \ldots Vn\}. B = \{bj(k)\}$, which represents the probability of harmony between the observed state and the hidden state. From the melody trend of the previous step, calculate the possible harmony hidden state:

$$Bj\ (k) = P(Ot = vk\ |qt = \theta j) \tag{7}$$

(3) Initial state probability matrix $\pi, \pi = \{i\}$, and:

$$\pi_i = P(qt = \theta j), 1 \leq \mathrm{i},\ \mathrm{j} \leq M \tag{8}$$

Most of the initial chords of popular music start from the tonic key, and the probability refers to the matrix in the database.

(4) Hidden state transition probability matrix $A = \{aij\}$.

$$aij = P(qt + 1 = \theta j\ |qt = \theta i) \tag{9}$$

Where $aij \gg 0$. State at the previous moment t is qt, and the chord $q(t + 1)$ is at the next time "$t + 1$", the Markov chain chord state transition as shown in Figure 4.

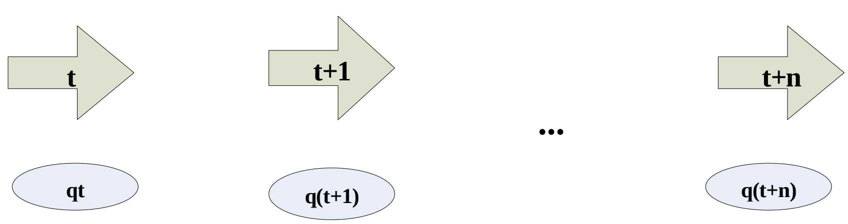


Figure 4. The Markov Chain Chord State Transition.

Accompaniment Matching Analysis Notes and harmony parts matching

In this paper, in terms of algorithmic programming, first, select possible chords for the melody, and then provide the probability of the current chord according to the previous chord state. In the melody limited chorus option, generate the required basic harmony and proceed to the next step. In the melody limited chorus option, generate the required basic harmony and proceed to the next step. Hidden Markov chains are used in basic chord generation. The latter includes probability generation ofs altered chords and suspended chords. The above chords can be obtained from the dataset. The harmony generation system flow chart is as illustrated in Figure 5.

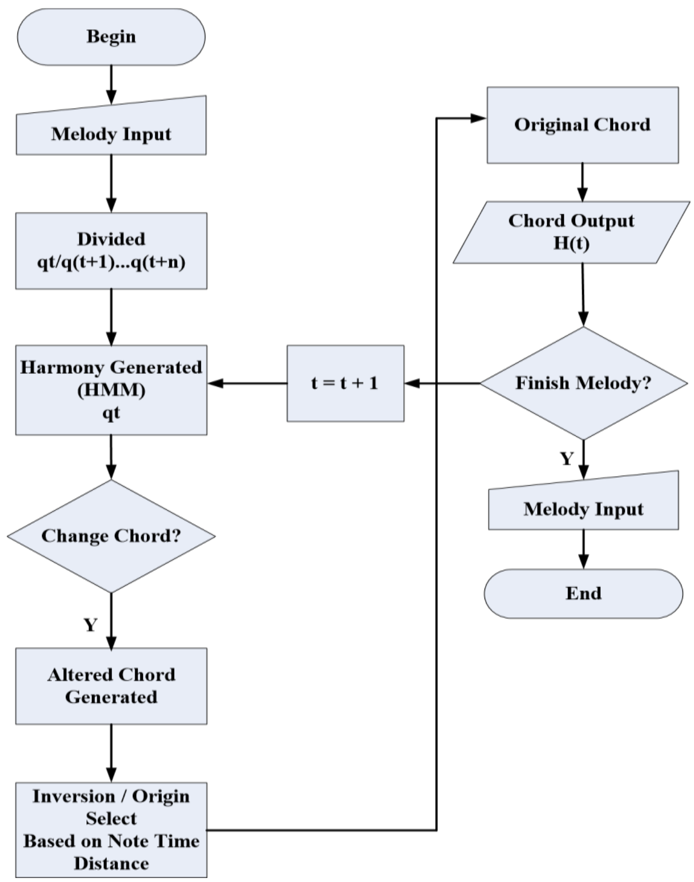


Figure 5. Harmony Generation Flow Chart.

For harmony texture part matching, keyboard and guitar can perform in a complementary role for both. When the keyboard performs a long tone for the harmony texture, the guitar can perform a rhythmic tone for the harmony texture, vice versa.

Chord progression example: I7 - VIm7 – I – IV - VIm7 – V - V7 – I. Keyboard performs a long tone as the background, and Arpeggio part can be performed by guitar, I7 chord including “C, E, G, B,” as the four parameters: 12k, 12k+4, 12k+7, 12k+11. Bass part matching uses original, and the second pitches from high to low are 12k, 12k+4, 12k+7, and 12k+11. And the pitch 12k+7, from low to light is 12+11, 12+12, 12+16, 12+12, 12+11, and 12+12.

Bass and percussion parts matching

Bass and percussion are representatives of the rhythm section of pop music. The bass part includes conventional bass, melody bass, and figured bass. The genetic algorithm can automatically compose music when provided with a bass line (De Prisco et al., 2010). According to previous research, the genetic algorithm can automatically generate the required bass-part according to the ready-made melody and chord progression direction data.

(1) Conventional Bass: according to the dataset to generate a conventional bass line.

(2) Melody Bass: decreasing big jump and increasing stepwise.

(3) Figured Bass: repeating pitches to avoid exceeding the range.

The example chord progression: I7 – VI7 – I – IV – VI7 – V – V7 – I. In the case of I7 – VI7, the first note is the root note of the original, and C can be connected with one of A, C, E, G notes from VI7. If change these notes to digits, 0 can be connected with one of 9, 12, 16, 18. Melodyized harmony uses the database to generate regular rhythms, so that the pitch goes up and down, not exceeding the range. After the above algorithm determines the pitch of the accent position, other parts use more passing tones.

The percussion section of pop music exhibits strong patternicity and belongs to the arrangement of the looping type. In the looping rhythm pattern, adding some drum group variations in each section can enhance the music’s agility. Therefore, in every 3-4 cycles, slight changes can be made on the basis of the original rhythm pattern to meet the changing needs. A piece of music typically changes its sectional texture in different parts, and the drum group rhythm pattern follows suit. The drum group generation process rules are shown in Figure 8.

(1) Drum Set Loop: with drum set loop rhythm.

(2) Drum Set Subtraction: in the basic drum set to delete separate strong beats.

(3) Drum Set Fine Tuning: in the basic drum set to change the note slightly.

(4) Embellishment Ending: in the ending part to add some embellishment notes.

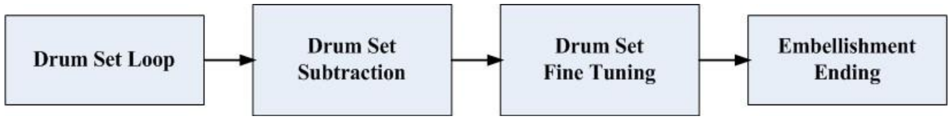


Figure 6. The Rules of Drum Set Generation Flow Chart.

In summary, this section centers around the architecture approach of the AI Score Following System based on the HMM. Starting from the system and database construction process, it provides a detailed introduction to the principles of setting up the two classes and three probability matrices of HMM. Furthermore, it explains the accompaniment matching methods for notes, harmony, and rhythm, emphasizing melody, texture, bass, and percussion sections. It can be understood as a mapping construction scheme for the AI Composition System based on HMM and music theory databases.

DISSCUSSION

The following experiment compares the original accompaniment of the three pieces: Love Transferred (Eason Chan, 2013), Rice Field (Jay Chou, 2013), and Glory (TFBOYS, 2021) with the proposed HMM-AICS in this article as an example. Table 8 shows the basic music attributes.

Pieces	Tempo	Key
Love Transferred	54	F Major
Rice Field	82	A Major
Glory	62	A minor

Table 8. The Basic Music Attributes.

Tables 9, 10, and 11 show the chord generation cases of the original chords, music memos, and HMM-AICS of three pieces of music.

Piece	Chord Direction				
	Original Chord Direction (1 Grid 1 Measure)				
Love Transferred	I	IV	IIIm	VIm – V	I
	VIm	IIm	IIm7	V	VIm
	III	IV – V	I	VIm	IIm – IIm7
	IIm7	VI	Vsus4add9	I	

	Music Memos				
	V	V – I	VI – IIIm	I	I
	I – V	I	IIm	I	VIm
	VIm	IV – V	I	I	I
	VIm	I	IIm – V	I	
	HMM-AICS				
	IIIm	II	V7	VIm – V	VIm
	Vsus2	IIm	IIm7	V	IIIm
	III7omit5	II	IIIm	I	IIIsus2
	V	II7	Vsus4	I	

Table 9. Love Transferred.

Piece	**Chord Direction**				
Rice Field	**Original Chord Direction (1 Grid 1 Measure)**				
	Iadd9 – V	VI7 – III7	II7 – Vsus4	Iadd9	Iadd9 – V
	VI7 – III7	IIm7 – Vsus4	Iadd9	Iadd9 – V	VI7 – III7
	IIm7 – Vsus4	Iadd9			
	Music Memos				
	I – IIIm	V	IIIm	VIm – V	IIIm
	VI7 – III7	VI	IIIM7 – II	V	V – IIm
	V – IV	I			
	HMM-AICS				
	III – II	II – I	II – V7	III	II
	III	IIm7sus2 – V	I	V7	VI – I
	IIm	II7	I		

Table 10. Rice Field.

Piece	**Chord Direction**			
Glory	**Original Chord Direction (1 Grid 1 Measure)**			
	Vm7 – I	VIm7 – VII7	IIIm7omit5 – VIM7	II_m – II
	Vm7 – I	VIm7 – VII7	IIIm7omit5 – VIM7	II_m
	Music Memos			
	Vm – IIm	#VII – II_m	I – VI_{m7}	#VII
	II_m – #VII	III_m – VI_m	V_m – VI_m	II_m
	HMM-AICS			
	I	Vm7omit3 – II7	IV7 – IIm7	I
	I – III_m	Vm7	bIII_m – VI_7	VII

Table 11. Glory.

We also conducted a comparative analysis of the musical content of the three versions of the aforementioned three pieces：

(1) About the Love Transferred

a. Original: mainly uses MIDI keyboard, strings, and Jazz drum set. Keyboard includes Arpeggio, strings secondary melody, and percussion is more regular.

b. Music Memos: uses bass, Jazz 2-part drum set, and drum kit accompaniment for the bass part, with a fixed form more rigid.

c. HMM-AICS: uses MIDI keyboard, guitar, and bass to perform the polyphonic accompaniment. Keyboard uses column chords, guitar accompaniment, and the bass uses a long bass tone.

(2) About the Rice Field

a. Original: mainly uses drum set, percussion, vocal choir, and synthesizer accompaniment. The abundant percussion timbres make the performance lively. The synthesizer in addition to arpeggio uses a bright high-frequency timbre.

b. Music Memos: uses bass and drum set, with some texture variance.

c. HMM-AICS: uses drum set, synthesizer, and guitar for the arrangement. Guitar performs arpeggio, drum set fixed rhythm is more rigid, and synthesizer supplements guitar texture and low-pitched parts.

(3) About the Glory

a. Original: uses drum set, percussion, vocal choir, synthesizer accompaniment. The arrangement is more regular, percussion uses loops texture, vocal choir uses choir harmony, and synthesizer uses high-pitched plucked for secondary melody.

b. Music Memos: uses bass and drum set. Texture has some variance.

c. HMM-AICS: uses drum set, synthesizer, guitar, and bass. Synthesizer performs plucked arpeggio, drum set performs looped texture with approximate variance. Bass performs low-pitched tones, and the guitar uses variant arpeggio for the supplement.

Based on the comparison results above, the proposed HMM-AICS in this paper can essentially accomplish the conventional arrangement of popular music styles. However, it is also evident that there is still considerable room for improvement with HMM-AICS. The accompaniment generated based on the melody in this study may benefit from further optimization using alternative algorithms, particularly in aspects such as harmony and texture.

CONCLUSION

This study developed an Artificial Intelligence Composition System (AICS) based on Hidden Markov Models (HMM) and a music theory database, achieving certain results in the generation of accompaniment music. The discussion on the construction of the music theory database revolves around the three databases: "Notes, Harmony, and Rhythm," with a particular focus on defining algorithms for musical elements such as pitch, rhythm, and duration. The architecture of HMM-AICS is explained starting from the process of system and database construction, providing a detailed introduction to the principles of setting up two classes and three probability matrices within HMM, and further explaining the methods for matching notes, harmony, and rhythm in accompaniment. HMM-AICS performed well in experiments and can serve as a case study for AI music generation for researchers.

Recommendations

This study belongs to the interdisciplinary research content of computer science and musicology, and can provide reference cases for researchers in these two disciplines. This study simulates the complete orchestration process in popular music and generates music through hidden Markov chains through machine learning. The main focus of this study that can provide reference for relevant researchers is as follows.

(1) The Architecture Method of HMM-AICS.

(2) Establish a music theory database - note database, harmony database, and rhythm database.

(3) The Principle of Four Step Generation Setting for Hidden Markov Chains

(4) Matching Method between HMM and Harmony Sound Part

(5) The mapping and matching method of accompaniment parts - notes, harmonies, bass, and Percussion parts.

From the perspective of accompaniment assistance, the proposed HMM-AICS still has significant room for development. AI Composition belongs to the challenging field of artificial intelligence, requiring interdisciplinary integration of artificial intelligence, machine learning, and music theory to achieve certain research results. This study references the complete process of music composition in an AI manner and then imitates and generates accompaniment styles, providing insights into new arranging and music generation methods. The proposed HMM-AICS is still in a relatively new stage. Due to various limitations, there are still many shortcomings. The fixed nature of accompaniment texture, the need for database expansion, and the improvement of music rules are challenges that need to be addressed and are directions for future research.

ACKNOWLEDGEMENTS

The authors express their deep appreciation to Kyonggi University in South Korea for the invaluable support throughout this project. Special thanks to Teacher Yanyuan Gao from Zhangzhou City Vocational College in China for her technical support to this research. This study serves as important research support for Weijia Yang's application for his doctoral degree of Kyonggi University. All authors made contributions to this study. There are no conflicts of interest in this study.

REFERENCES

Allan, M., & Williams, C. (2004). Harmonising chorales by probabilistic inference. Advances in Neural Information Processing Systems, 17.

Bäckman, K. (2009). Automatic jazz harmony evolution. Proceedings of the 6th Sound and Music Computing Conference (SMC), 349–354.

Band-in-a-Box. (2024, Feb. 6). https://www.pgmusic.com

Bell, C. (2011). Algorithmic music composition using dynamic Markov chains and genetic algorithms. Journal of Computing Sciences in Colleges, 27(2), 99– 107.

Chong, E. K. M., & Ding, Q. (2014). Symbolic representation of chords for rulebased evaluation of tonal progressions.

Dannenberg, R. B., & Grubb, L. (1994). Automating ensemble performance. Procs of ICMC94, 63–69.

Dannenberg, R. B., & Hu, N. (2003). Polyphonic audio matching for score following and intelligent audio editors.

De Prisco, R., Zaccagnino, G., & Zaccagnino, R. (2010). Evobasscomposer: A multiobjective genetic algorithm for 4-voice compositions. Proceedings of the 12th Annual Conference on Genetic and Evolutionary Computation, 817–818.

Doush, I. A., & Sawalha, A. (2020). Automatic music composition using genetic algorithm and artificial neural networks. Malaysian Journal of Computer Science, 33(1), 35–51.

Eason Chan. (2013). Love Transfer. https://youtu.be/jouvNcQOm7Q?si=X5ELaq2TI76bNxO8

Farbood, M., & Schöner, B. (2001). Analysis and synthesis of Palestrina-style counterpoint using Markov chains. ICMC.

Fernández, J. D., & Vico, F. (2013). AI methods in algorithmic composition: A comprehensive survey. Journal of Artificial Intelligence Research, 48, 513–582.

Gurkaynak, G., Yilmaz, I., & Haksever, G. (2016). Stifling artificial intelligence: Human perils. Computer Law & Security Review, 32(5), 749–758.

Herremans, D., Chuan, C.-H., & Chew, E. (2017). A functional taxonomy of music generation systems. ACM Computing Surveys (CSUR), 50(5), 1–30.

Jay Chou. (2013). Rice Field. https://youtu.be/kC2bkbMC6Zs?si=kNk971D1PEseDxUe

Jordanous, A., & Smaill, A. (2008). Artificially intelligent accompaniment using Hidden Markov Models to model musical structure.

Linson, A., Dobbyn, C., & Laney, R. (2012). Improvisation without representation: Artificial intelligence and music.

Luque, S. (2009). The stochastic synthesis of iannis xenakis. Leonardo Music Journal, 19, 77–84.

Minsky, M. (2007). The emotion machine: Commonsense thinking, artificial intelligence, and the future of the human mind. Simon and Schuster.

Music Memos. (2024, Feb. 6). https://www.chordai.net/musicmemos/

Oliwa, T. M. (2008). Genetic algorithms and the abc music notation language for rock music composition. Proceedings of the 10th Annual Conference on Genetic and Evolutionary Computation, 1603–1610.

Orio, N. (2001). An automatic accompanist based on hidden markov models. Congress of the Italian Association for Artificial Intelligence, 64–69.

Rohrmeier, M., & Graepel, T. (2012). Comparing feature-based models of harmony. Proceedings of the 9th International Symposium on Computer Music Modelling and Retrieval, 357–370.

Sako, S., Yamamoto, R., & Kitamura, T. (2014). Ryry: A real-time score-following automatic accompaniment playback system capable of real performances with errors, repeats and jumps. Active Media Technology: 10th International Conference, AMT 2014, Warsaw, Poland, August 11-14, 2014. Proceedings 10, 134–145.

Shan, M., & Chiu, S. (2010). Algorithmic compositions based on discovered musical patterns. MULTIMEDIA TOOLS AND APPLICATIONS, 46(1), 1–23.

Simon, I., Morris, D., & Basu, S. (2008). MySong: Automatic accompaniment generation for vocal melodies. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 725–734.

Tang, F. (2021). Analysis on the construction of english language and culture teaching system based on artificial intelligence. Big Data Analytics for Cyber-Physical System in Smart City: BDCPS 2020, 28-29 December 2020, Shanghai, China, 1464–1469.

TFBOYS. (2021). Glory. https://youtu.be/8CookdwavLo?si=NQE9REIBPMMXAmOA Thom, B. (2000). Artificial intelligence and real-time interactive improvisation. Proceedings from the AAAI-2000 Music and AI Workshop, 35–39.

Wright, R., Younker, B. A., Beynon, C., Hutchison, J., Linton, L., Beynon, S., Davidson, B., & Duarte, N. (2012). Tuning into the future: Sharing initial insights about the 2012 Musical Futures pilot project in Ontario. The Canadian Music Educator, 53(4), 14.

Xia, G., & Dannenberg, R. B. (2015). Duet interaction: Learning musicianship for automatic accompaniment. NIME, 259–264.

Zarro, R. D., & Anwer, M. A. (2017). Recognition-based online Kurdish character recognition using hidden Markov model and harmony search. Engineering Science and Technology, an International Journal, 20(2), 783–794.

Zhu, D. (2017). Analysis of the application of artificial intelligence in college English teaching. 2017 2nd International Conference on Control, Automation and Artificial Intelligence (CAAI 2017), 235–237.

GENİŞLETİLMİŞ ÖZET

Bu çalışma, Gizli Markov Modelleri (HMM) ve bir müzik teorisi veritabanına dayalı bir YZ Besteleme Sistemi (AICS) geliştirmiş ve eşlik müziği üretiminde belirli sonuçlar elde etmiştir. Müzik teorisi veritabanının inşası, “Nota, Armoni ve Ritim” olmak üzere üç veritabanı etrafında açıklanmış, özellikle perde, ritim ve süre gibi müzikal unsurların algoritmalarla tanımlanmasına odaklanılmıştır. HMM-AICS’nin mimarisi, sistem ve veritabanı oluşturma sürecinden başlayarak, HMM içindeki iki sınıf ve üç olasılık matrisinin ayar prensiplerine ayrıntılı bir giriş sağlamış ve ardından nota, armoni ve ritmin eşlikte nasıl eşleştirileceğini açıklamıştır. HMM-AICS, deneylerde iyi performans sergilemiş ve AI müzik üretimi için ilgili araştırmacılara bir vaka çalışması olarak hizmet edebilir.

HMM-AICS, popüler müzik tarzlarının müzik teorisi ve orkestrasyonunu bir araya getirerek, teori tabanlı bir veritabanı inşasını tamamlamış ve bu sisteme Gizli Markov Modellerini entegre etmiştir. HMM-AICS veritabanı müzik teorisine dayanır ve bu sayede eşlik, müzik kompozisyonunun kurallarına daha uygun hale getirilir ve programın işleyişi optimize edilir. Bu inşa, "Nota Armoni ve Ritim" olmak üzere üç veritabanı etrafında gerçekleşir. Nota veritabanı, müziğin tınısı, perde ve süre gibi temel verileri sağlar. Tonal veritabanı, armoni veya melodinin majör ve minör skalalarına ve diğer özel tonal yönlere odaklanır. Armoni veritabanı, otomatik eşlik için temel bir veri kaynağıdır ve yaygın üçlü akorlar, yedili akorlar, askı akorları vb. içerir. Ayrıca, rastgele olasılıkla ilgili algoritmalarla birleştirerek armonik kelime dağarcığının zenginliğini artırabilir. Ritim veritabanı, müziğin düşük sesli bölümlerine referans olarak hizmet eder ve bas ve davul bölümlerini içerir. Nota sürelerine ve değerlerine göre müzik ritmi yazılır, birkaç ritim gruplanır ve ritim veritabanında saklanır. Dinamikler ve tını bölümü için geçici olarak MIDI CC verileri kullanılabilir ve daha sonra üçüncü parti VST eklentileriyle ses kalitesi genişletilebilir. Veritabanı, müzikal unsurların analizi, armoni eşleştirme, orkestrasyon eşleştirme gibi bölümlerde kullanılabilir.

HMM-AICS'nin çekirdeği veritabanı inşasıdır ve mimari haritası, HMM içindeki iki sınıf ve üç olasılık matrisinin ayarına odaklanır. Nota, armoni ve ritim eşlik yöntemleri, HMM'nin ayırt edici bir özelliğidir ve özellikle melodi, doku, bas ve vurmalı çalgılar bölümlerine vurgu yapar. HMM-AICS, MIDI dosya verilerine dayanarak otomatik olarak akorlar ve eşlik dokuları oluşturabilir. Armoni düzenleme süreci, akor eşleştirme ve armoni eşleştirme bölümlerini içeren eşlik üretim veritabanı uygulama sürecinin bir parçasıdır. Akor eşleştirme bölümü, temel akor eşleştirme, değiştirilmiş akor eşleştirme ve ters çevrim/kök pozisyonundan oluşur. Armoni eşleştirme bölümü ise vurmalı çalgılar bölümü eşleştirme, bas bölümü eşleştirme, armoni dokusu eşleştirme ve ikincil melodi bölümü eşleştirmeden oluşur. Gizli Markov algoritması, bir sonraki armoni veya akoru tahmin etmek için kullanılır. Besteleme sürecinin standardizasyon eğilimini, algoritmik kompozisyon sistemini takip eden bir nota veya "Band-in-a-Box" ve "Music Memos" gibi klasik araçlar temsil eder.

Bu çalışmanın son bölümünde, müzik eserlerinin deneysel doğrulaması da yapılmıştır. Deneyler, bu çalışmada önerilen HMM-AICS'yi bir üretim modeli olarak kullanmıştır. Eason Chan'ın 2013 yılında yayınlanan "Love Transfer", Jay Chou'nun 2013 yılında yayınlanan "Rice Field" ve TFBOYS'un 2021 yılında yayınlanan "Glory" adlı üç şarkısının orijinal eşlikleriyle karşılaştırmalar yapılmıştır. Müzik içeriğinin niteliksel araştırma sonuçlarına dayanan karşılaştırmalara göre, önerilen HMM-AICS temel olarak pop müzik tarzının rutin düzenlemelerini gerçekleştirebilmektedir. Bununla birlikte, HMM-AICS'nin hala önemli bir ilerleme kaydetme potansiyeline sahip olduğu da gözlemlenmiştir. Bu çalışmada melodiye dayalı olarak üretilen eşliklerde, armoni, doku gibi detaylarda diğer algoritmaların kullanılarak daha fazla optimize edilmesi düşünülebilir.

HMM-AICS, deneysel ortamlarda güçlü bir performans sergilemiş ve YZ müzik üretimi konusunda araştırmacılar için değerli bir vaka çalışması olarak konumlandırılmıştır. Bu disiplinler arası çalışma, bilgisayar bilimi ve müzikoloji arasındaki köprüyü kurarak her iki alandaki araştırmacılara referans çerçevesi sunmaktadır. Çalışmanın ana katkıları şunlardır:

(1) HMM-AICS'nin mimari metodolojisi.

(2) Nota, Armoni ve Ritim veritabanlarının oluşturulması ve yapılandırılması.

(3) Gizli Markov Zincirleri için dört adımlı üretim ayar ilkeleri.

(4) HMM ile armonik ses bileşenlerinin eşleştirilmesi yöntemleri.

(5) Eşlik unsurları olan nota, armoni, bas ve perküsyonun haritalanması ve eşleştirilmesi stratejileri.

HMM-AICS modelinin, eşlik müziği yardımı alanında önemli bir gelişim potansiyeline sahip olduğu düşünülebilir. YZ destekli besteleme, birçok disiplinle yakından ilişkili bir çapraz disiplin alanıdır ve bu durum, HMM-AICS'nin bir disiplinler arası entegrasyon sonucu olarak gelişiminin de YZ, makine öğrenimi ve müzik teorisi gibi alanlarda derinlemesine işbirliği gerektirdiğini göstermektedir. Bu çalışma, zorlu bir alan olan YZ destekli bestelemeyi temel alarak, müzik yaratım sürecini referans almakta ve eşlik tarzlarının taklidi ve üretimi üzerine yoğunlaşmaktadır. Bu, müzik üretiminde yenilik açısından önemli bir anlam taşımakla birlikte, müzik üretim yöntemlerine de yeni bir yön sunmaktadır.

Ancak, HMM-AICS şu anda bazı sınırlamalar ve zorluklarla karşı karşıyadır. Sistem, müzik içeriğinin yalnızca armoni kısmına odaklanmakta ve tını, kontrpuan gibi alanlarda daha fazla tartışmaya yer vermemektedir. Özellikle, eşlik dokusunun nispeten sabit kalması sorunu, diğer algoritmalarla çözülmesi gereken bir durum olabilir ve güncellemelerle daha zengin ve esnek müzikal ifadeler sağlanabilir. Ayrıca, veritabanının zenginliği, YZ öğrenme sürecini doğrudan etkiler; daha kapsamlı bir müzik veritabanı, HMM-AICS'nin performansını artırmaya yardımcı olacaktır. Aynı zamanda, müzik teorisinin bütünlüğü de kilit bir faktördür ve daha doğru müzik yaratımı ve eşlik üretimi için müzik teorisi ile Aapay Zeka algoritmalarının daha derin bir şekilde birleştirilmesi gerekmektedir. HMM-AICS'nin gelecekteki araştırmalarda, algoritma optimizasyonu, veritabanı genişletme ve müzik teorisinin derinleştirilmesi gibi konularda daha fazla geliştirilmesi gerekecektir. Sürekli keşif ve yenilikler sayesinde, HMM-AICS'nin YZ destekli müzik yaratımı alanında daha büyük atılımlar yapması beklenebilir.