Volume: 2 | Issue: 10 | Received-24-09-2025 | Accepted-29-09-2025 | Published-06-10-2025

UAI JOURNAL OF ARTS, HUMANITIES AND SOCIAL SCIENCES (UAIJAHSS)

Abbreviated Key Title: UAI J Arts Humanit Soc Sci
ISSN: 3048-7692 (Online)
Journal Homepage: https://uaipublisher.com/uaijahss/
Volume- 2 Issue- 10 (October) 2025
Frequency: Monthly

Visualization Research of Vocal Music Works——A Case Study on the Song "Phoenix Tree"

JIA LI Ariunaa Sh.

Graduate University of Mongolia, Ulaanbaatar Mongolia.

Corresponding Author: JIA LI Ariunaa Sh.

ABSTRACT

Current issues in vocal performance research primarily revolve around the insufficient utilization of resources such as sheet music, audio-visual materials, and practical performances, as well as the failure to integrate modern empirical methods with traditional vocal research approaches. This study focuses on using the Vmus music visualization analysis program to analyze elements such as rhythm, tempo, dynamics, and timbre in the simple folk song "Phoenix Tree". By employing measurement tools to compare different performance versions, it demonstrates how to achieve better interpretations of musical works.

KEY WORDS: Visualization, Vocal music, Vocal performance, Worm plot

Introduction

In recent years, music visualization technology has gradually filled gaps in traditional art research. Through audio spectrum analysis, Motion Capture systems, and AI algorithms, researchers can transform subjective musical aesthetics into quantifiable data. For example, using Sonic Visualizer to generate spectrograms of opera performances allows for precise analysis of the correlation between a singer's vibrato frequency (typically 4-6Hz) and emotional intensity. Motion Capture technology, which tracks the movement trajectories of instrumental performers, can establish mathematical models (e.g., the functional relationship between a pianist's keystroke angle and tonal brightness) linking performance gestures to musical tension. This empirical shift has given rise to the new field of "Digital Musicology." A Cambridge University team employed machine learning to analyze 6,000 historical recordings, reconstructing the evolutionary path of 20th-century vocal techniques. Their findings revealed that the transition from chest to abdominal breathing support was not linear but exhibited regional discontinuities. This study introduces several current music visualization software tools and recent academic achievements, with a focus on the empirical analysis of the Chinese vocal piece "Phoenix Tree" using the Vmus music visualization analysis software developed by Professor Yang Jian's team at the Shanghai Conservatory of Music.

1. Types of Music Visualization Software

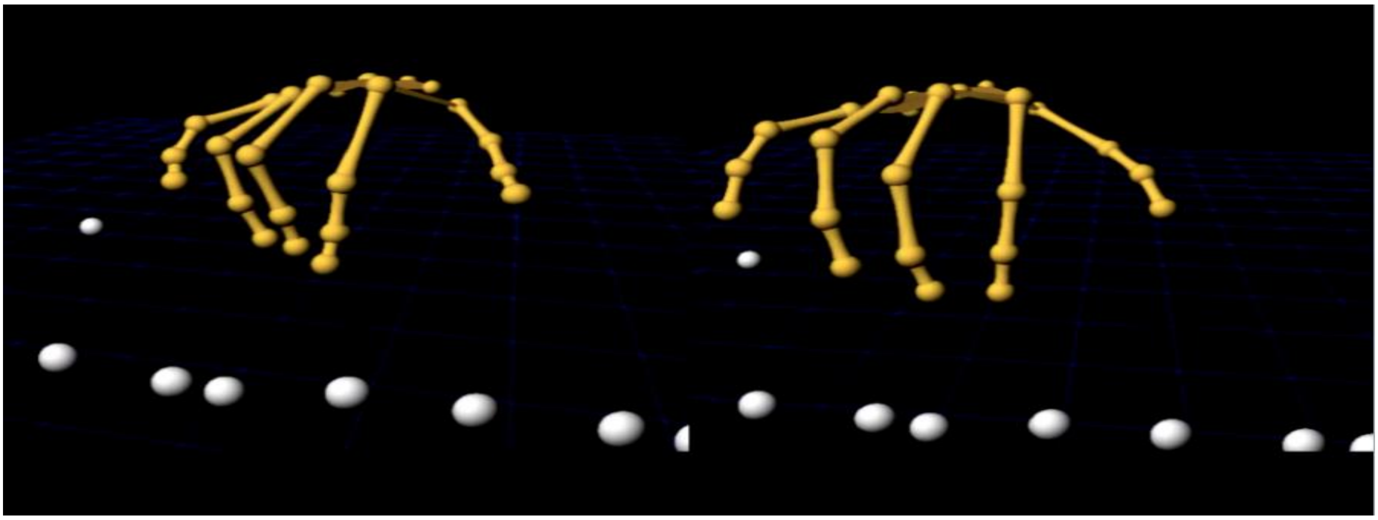
Currently, the most widely used music visualization software internationally can be categorized into real-time music visualization software (suitable for live performances/interactive scenarios), video editing-integrated tools (for music videos/social media content creation), mobile applications (Android/iOS), open-source and professional analysis tools (for research/in-depth analysis), and online platforms & AI tools.

Real-time music visualization software (suitable for live performances/interactive scenarios): 1) HeavyM – Supports real-time

DOI: 10.5281/zenodo.17277096

audio synchronization and projection mapping technology, transforming any surface into a dynamic visual interface. It is ideal for large-scale events like music festivals and art installations. Its advantages include a user-friendly interface, customizable templates, and multi-screen output support; 2) Synesthesia Live – Features a modular design, allowing users to customize visual effects through a node-based interface. It supports MIDI controller real-time manipulation, making it suitable for electronic music performances. Its standout feature is real-time audio analysis, generating abstract graphics that dynamically match the music; 3) Magic Music Visuals – Provides real-time audio input response, compatible with Windows/Mac. It includes various dynamic templates (e.g., particle effects, spectral waveforms), making it ideal for DJs and VJs for improvisational creation; and 4) Motion Capture-Tracks the limb movement trajectories of instrumental performers, enabling the construction of mathematical models that link performance gestures to musical tension (e.g., the functional relationship between a pianist's keystroke angle and tonal brightness).

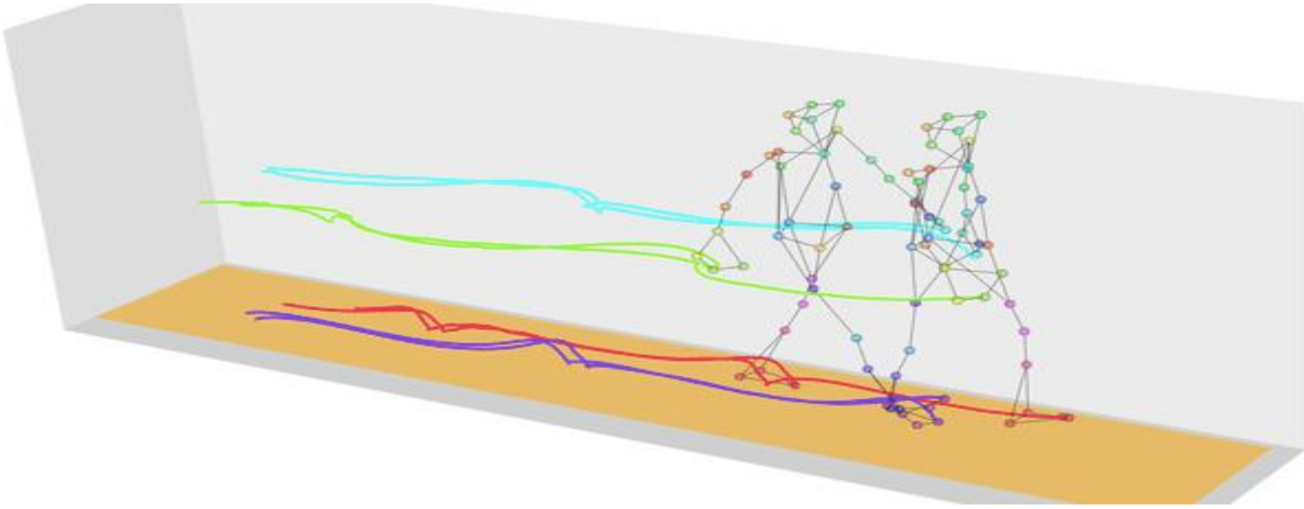
Figure 1. *Comparison of Performance Movements Captured by Motion Capture from Two Musicians*



Source: *https://en.wikipedia.org/wiki/Motion_capture* .

The above is the same musical passages performed by two pianists, presented in slow motion. This allows for a clear observation of the distinct body language expressions of the performers during their interpretation of the piece.

Figure 2. *Comparison of performance movements captured by Motion Capture from two* musicians.



Source: *https://en.wikipedia.org/wiki/Motion_capture*.

This technology allows for comparative analysis of each singer's body language when performing the same piece, enabling better artistic interpretation. However, it requires high technical proficiency, and the Motion Capture needs to fit closely to the singer's body, which may affect their normal vocal performance.

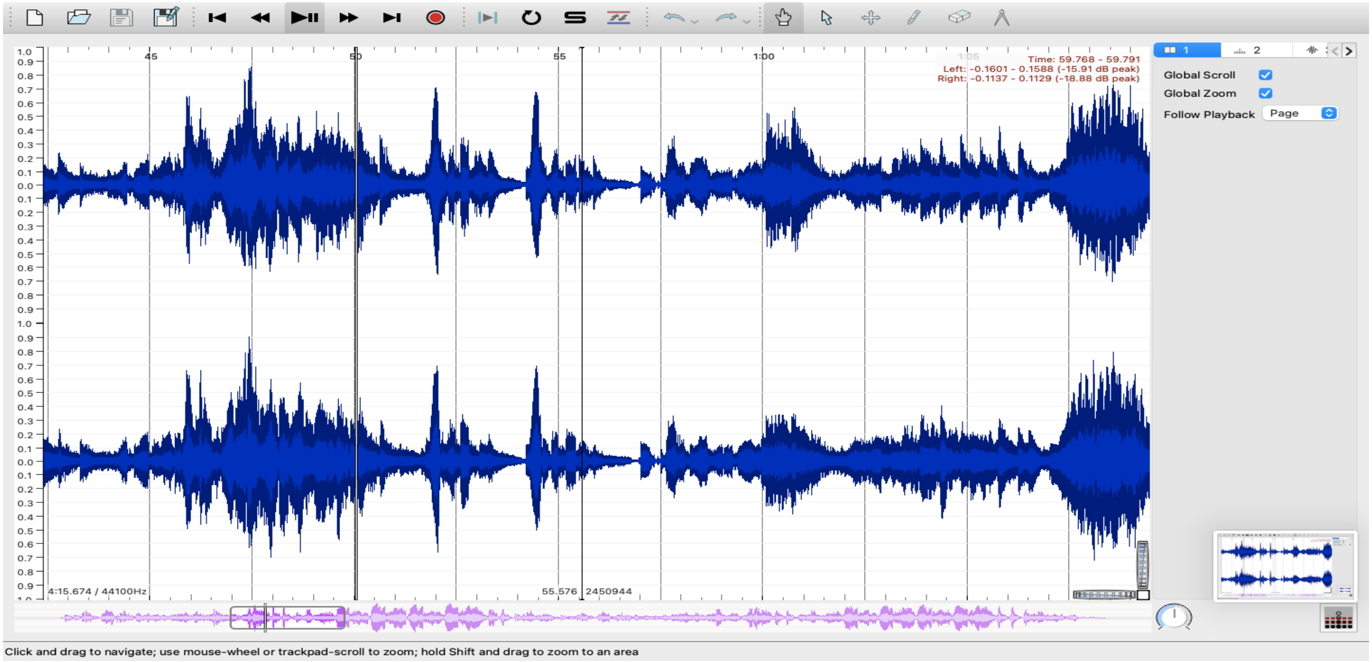
Video editing-integrated tools (suitable for music videos/social media content creation) include: 1) Adobe After Effects–A professional-grade tool that achieves high-precision visualization through audio spectrum and waveform effect plugins. It supports 3D layers and motion graphics, making it ideal for film/TV-level production. For example, its “Audio Spectrum” plugin generates beat-synchronized animations; 2) Veed–An online tool offering soundwave templates and automatic beat-sync functionality, with support for one-click HD video exports. It is well-suited for quick social media short video production. A key advantage is its built-in royalty-free music library, enabling direct generation of branded visual content; and 3) CapCut (International Version of Jianying)–Features integrated audio spectrum visualization, along with noise reduction and voice enhancement, making it a practical choice for musicians creating lyric videos or promotional clips.

DOI: 10.5281/zenodo.17277096

Mobile visualization apps (Android/iOS) encompass: 1) Avee Music Player (Android) – A highly customizable visualization tool that supports color/shape adjustments and HD video exports, ideal for creating personalized music videos; 2) Muviz Edge (Android) – Displays real-time spectrum effects on screen edges, optimized for full-screen devices. It integrates with third-party apps like Spotify and has low system resource usage. 3) Lumen (iOS) – A real-time visual synthesizer for mobile, featuring MIDI integration and LED screen compatibility, suitable for small live performances or electronic music production.

Open-source & professional analysis tools (for research/in-depth analysis) include: 1) Sonic Visualiser – An open-source tool supporting spectrograms, pitch tracking, and multi-layer annotations. It is widely used in music structure analysis and speech research, making it highly suitable for academic studies.

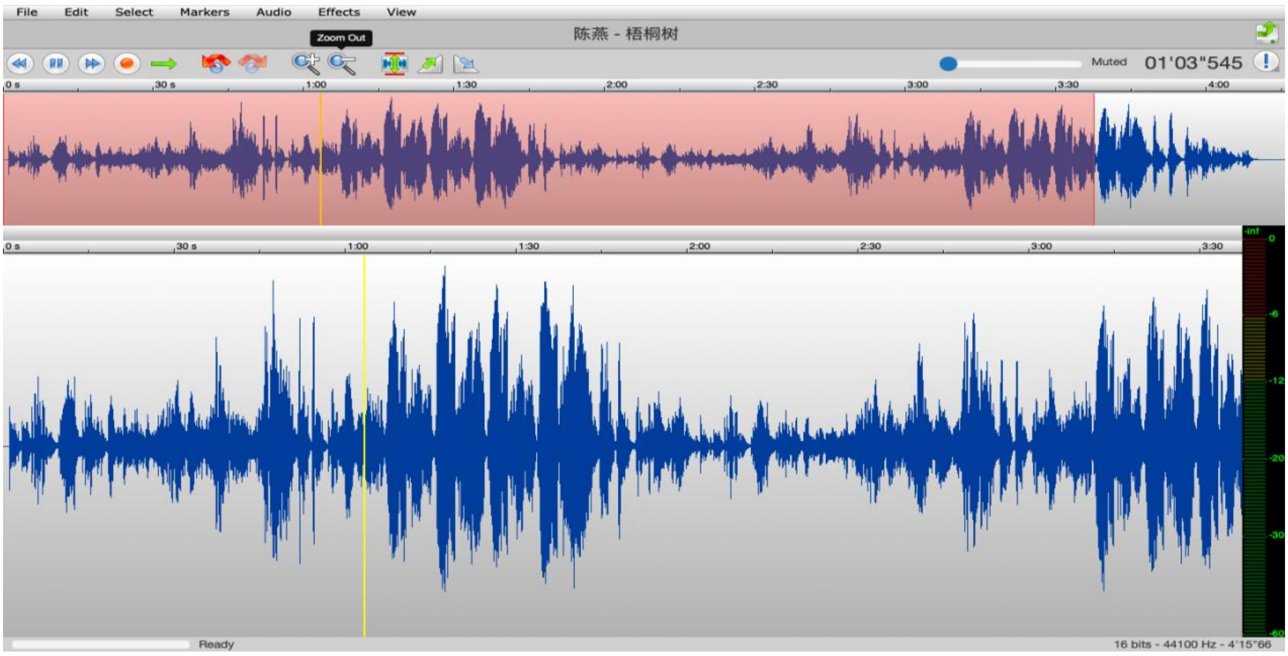
Figure 3. *Sonic Visualiser software interface.*



Source: *Sonic Visualiser APP*

2) Python + LibROSA – Combines Python libraries for audio feature extraction (e.g., beat tracking, Mel spectrograms), making it ideal for developing customized visualization algorithms; 3) Audacity – A free audio editor offering basic spectrum analysis functions, suitable for entry-level users performing simple audio visualization; 4) TwistedWave – A cross-platform audio editing tool compatible with Mac, Windows, iOS/iPadOS, and web browsers, renowned for its efficiency and multi-scenario adaptability. Its key advantages include real-time waveform display and instant undo/redo functionality, making it well-suited for educators demonstrating audio editing techniques.

Figure 4. *TwistedWave software interface.*



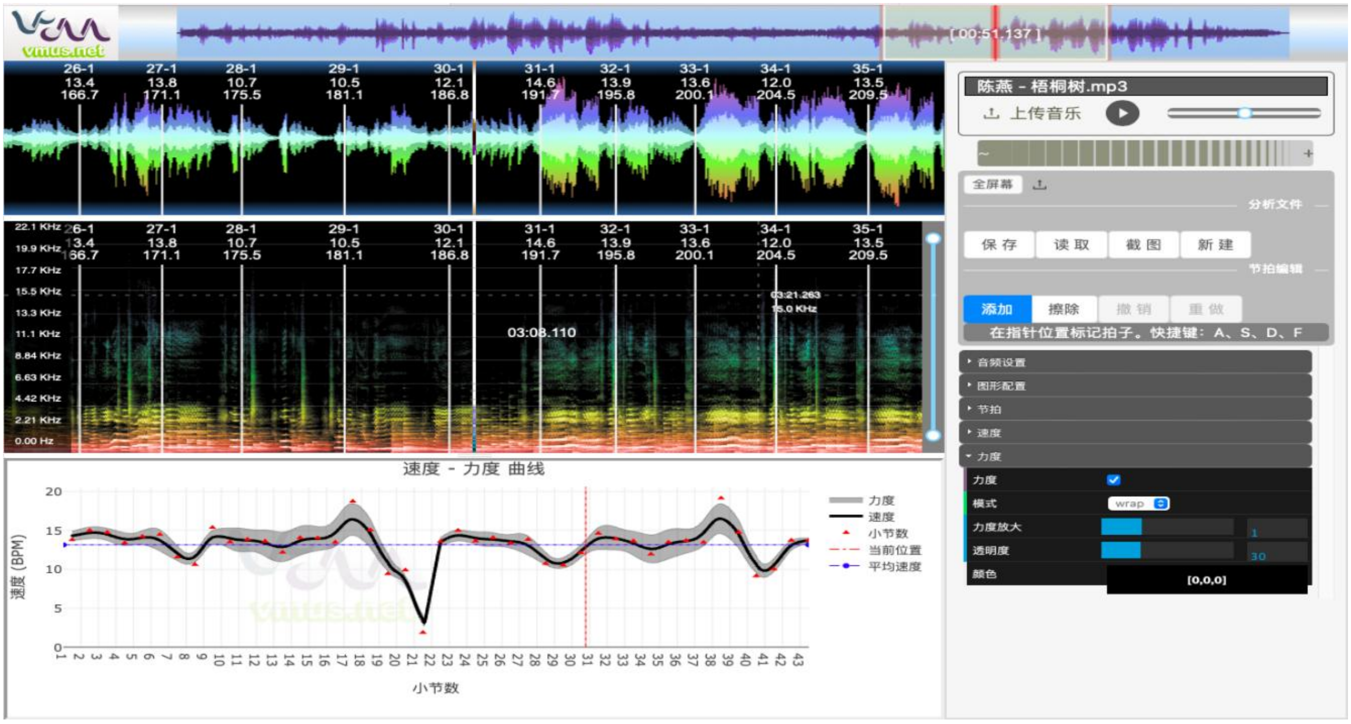
Source: *TwistedWave APP*

DOI: 10.5281/zenodo.17277096

Volume: 2 | Issue: 10 | Received-24-09-2025 | Accepted-29-09-2025 | Published-06-10-2025

5) Vmus.net. Developed by Professor Yang Jian's team at the Shanghai Conservatory of Music, Vmus.net is an online music visualization tool specializing in real-time performance analysis, with particular strengths in tempo, dynamics, and rhythm analysis. It generates intuitive tempo-dynamics curves that clearly display performance nuances such as crescendo/decrescendo and rubato. Its distinctive Performance Worm feature presents a three-dimensional dynamic model illustrating the interplay between tempo, dynamics and temporal progression, making it particularly effective for analyzing the accumulation and release of musical tension. The IOI deviation analysis quantitatively measures the discrepancy between actual note durations and theoretical values, revealing performers' individualized rhythmic treatments (e.g., elongation or compression of double-dotted rhythms). For instance, Maria Callas' dramatic tension-building techniques in *La Traviata* - including prolonged B-flat notes and manipulated double-dotted rhythms - can be efficiently analyzed through this platform. Notably, Vmus.net offers exceptional accessibility and compatibility as a browser-based tool requiring no installation. Users can simply drag-and-drop audio files for immediate analysis, with exportable visuals suitable for direct inclusion in academic papers or presentation materials. Its streamlined interface facilitates rapid visualization generation, proving especially valuable for real-time classroom demonstrations. However, the platform does present certain limitations: its analytical dimensions remain constrained, primarily focusing on temporal-domain parameters (tempo, dynamics, rhythm) while lacking advanced spectral analysis capabilities (such as formant tracking or harmonic structure decomposition). The system cannot process multi-track audio alignment or complex annotation requirements. Furthermore, compared to international counterparts like Sonic Visualiser, Vmus.net faces challenges in technical maintenance and updates due to relatively limited development funding and smaller team size, resulting in less frequent software enhancements.

Figure5. *Vmus.net software interface.*



Source: *Vmus.net*

Online platforms & AI tools include: Renderforest – Offers template-based music visualization generation with support for 3D scenes and animated text, ideal for quickly producing commercial promotional videos; 2) Syqel – AI-powered platform that generates dynamic visual effects in real-time, featuring cloud-updated templates suitable for users with limited design experience; and 3) Specterr – An online tool supporting particle effects and lyric synchronization, providing high-resolution, watermark-free exports perfect for independent musicians promoting their work.

While current software often proves overly complex, limiting the widespread adoption of such analytical methods in the music community, researchers and enthusiasts now have access to a simple yet powerful solution: the VMUS.net online music performance analysis platform.

However, these technological advancements also prompt reflection. AI-generated music may adhere to compositional rules but lacks the human warmth of historical context. Empirical music research enables precise quantitative analysis, aiding performers in technical execution, yet risks diminishing the art's expressive depth. This methodological evolution represents both an opportunity and a challenge for digital humanities in art research—enhancing analytical capabilities while testing the academic community's value judgments. Ultimately, this author advocates for a balanced approach: integrating traditional research methods with technological tools to achieve more holistic and meaningful interpretations of musical works.

2. Visualization-Based Music Research on the Art Song "Phoenix Tree"

2.1 Analysis of Musical Structure

"Phoenix Tree" is a lyrical Chinese art song composed by Xi Qiming with lyrics by Yang Zhanye. Originally performed by Chinese soprano Chen Yan, it was included in her 2010 album "In Search of Lost Elegance." This art song depicts autumn winds rustling through phoenix trees, using imagery of falling leaves and fruits as metaphors for the passage of time, expressing nostalgia for childhood and youth while contemplating the cycle of life.

DOI: 10.5281/zenodo.17277096

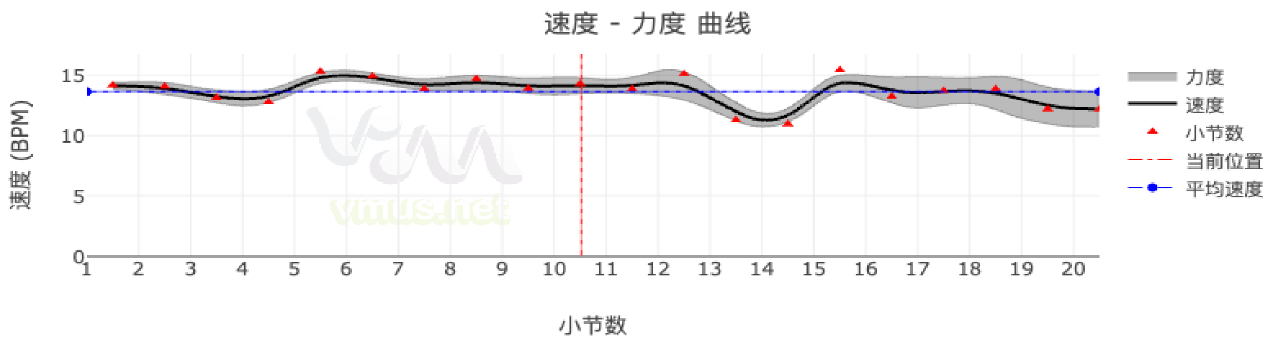
The piece is written in E-flat major with a 4/4 time signature, marked "Andantino" at approximately 76 beats per minute. Comprising 25 measures in total, it follows a compound binary form (A-B) that reflects the traditional musical narrative structure of "introduction-development-transition-conclusion."

2.2 Visualization Analysis of "Phoenix Tree"

Focusing on the visualization analysis of the art song "Phoenix Tree," this study employs three analytical tools from Vmus.net software: the Spectrogram, Music Worm Plot, and Tempo-Dynamics Plot for detailed examination. Below are the specific analytical methods for each visualization type, along with practical examples from "Phoenix Tree." Through quantitative analysis of vocal versions regarding pitch control, rhythmic treatment, and emotional expression, this approach aids in studying performance styles and affective interpretation.

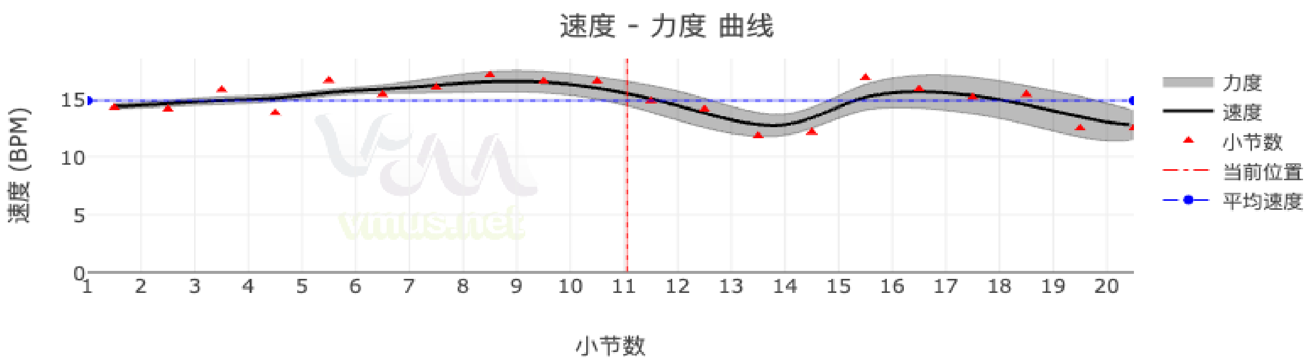
The following diagrams present a visual comparison of measures 1-19 (Section A) from the song "Phoenix Tree" performed by vocalists Chen Yan and Huang Lu. In this study, using the Tempo-Dynamics Plot, we model the relationship between instantaneous tempo variations (BPM) and relative dynamics (volume) in each measure, enabling quantitative comparison of the two performers' distinct approaches to rhythmic fluidity, phrasing tension, and lyrical expression.

Figure6. *Vmus.net CHEN YAN'Phoenix Tree'Tempo-Dynamics Plot*



Source: *Vmus.net*

Figure7. *Vmus.net HUANG LU'Phoenix Tree'Tempo-Dynamics Plot*



Source: *Vmus.net*

The black curve represents the trend of instantaneous tempo changes over time (by measure number); the gray shaded area indicates the dynamic range within each measure; the blue dashed line shows the average tempo of the entire section; the red dashed line marks the current comparison point; and the red triangle symbols denote measure starting points for rhythmic reference.

From the above comparison, it can be concluded that:

Dimension	CHEN YAN（Figure7）	HUANG LU（Figure7）
Average Tempo	Slightly slower (about 13.8 BPM)	Slightly faster (about 15.2 BPM)
Overall Fluctuation Range	More stable control, max fluctuation <4BPM	Higher fluctuation frequency, max amplitude reaching 6BPM
Rhythm Pattern	Parabolic structure: "acceleration in middle, deceleration at both ends"	Dramatic "fast-slow-fast" alternating pattern
Width of Gray Shaded Area	Overall narrower, fluctuating between approximately 10-14, progressing steadily, meeting traditional lyrical singing requirements.	Wider shaded areas, extreme amplitude in some measures (e.g., 9, 16), showing "trapezoidal peak-valley" alternation with strong emotional tension
Stylistic Approach	Structural integrity, rhythmic stability, and lyricism, leaning towards academic and technically rigorous style.	Expressive tension, rhythmic freedom, dramatic - suitable for stage performers/emotion-driven singers

Chen Yan adopts a tempo model more aligned with the "Largo cantabile" tradition, emphasizing natural transitions between phrases while pursuing coherence and depth. Huang Lu, however, favors a

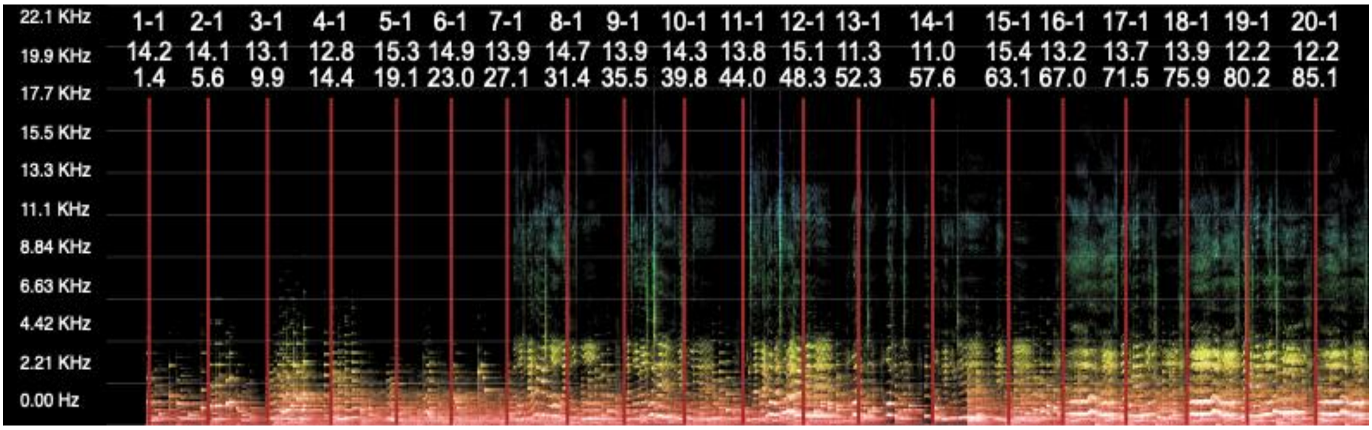
DOI: 10.5281/zenodo.17277096

"rubato" approach, incorporating significant deceleration at key emotional points (e.g., measures 9-11), demonstrating acute sensitivity to musical dramatic tension. In measures 14-17, Chen Yan maintains a moderate tempo with gradually restrained dynamics, creating a lingering resonance effect. Huang Lu, however, executes sudden accelerando coupled with intensified dynamics in measures 15-16, producing an emotional catharsis that transforms melancholic sentiment into what might be perceived as vocal outcry or liberation in the coda. Regarding dynamic control, Chen Yan maintains remarkable consistency, prioritizing phrasal continuity and vertical layering. Huang Lu employs a "sectional intensification with terminal release" strategy, utilizing dynamic fluctuations to amplify the music's expressiveness.

The tempo-dynamics plot visualization not only provides quantitative evidence for studying the art song "Phoenix Tree" but also reveals the underlying emotional logic, breath control, tempo organization, and dynamic expression strategies across different performance styles. This cross-dimensional interpretation from "visual music imagery" to "vocal expression style" assists singers in: identifying strengths and blind spots in their performance style; selecting appropriate rhythmic and dynamic strategies for the emotional core of the work; and achieving three-dimensional unity in "music-text-emotion" integration based on lyrical content.

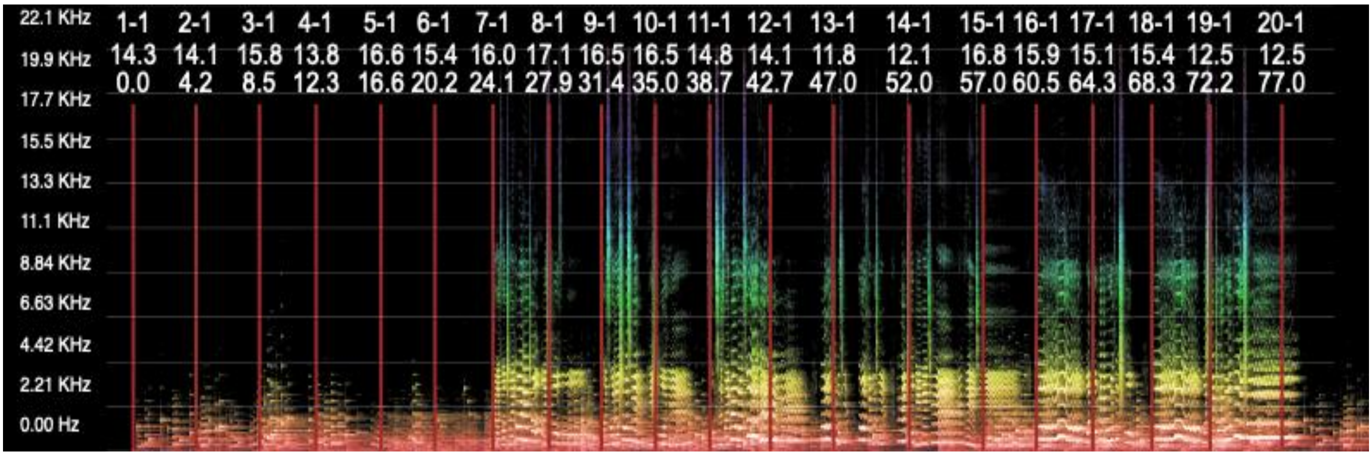
The following presents an in-depth, professional comparative analysis of the spectrograms of the two performers in the art song "Phoenix Tree." These images, based on frequency (vertical axis, Hz) and time (horizontal axis, measure divisions), use color intensity to represent sound strength and frequency components, revealing critical vocal characteristics such as pitch control, resonance usage, diction clarity, and emotional tension.

Figure*8. Vmus.net CHEN YAN'Phoenix Tree'Spectrogram*



Source: *Vmus.net*

Figure*9. Vmus.net HUANG LU'Phoenix Tree'Spectrogram*



Source: *Vmus.net*

High-frequency range (8.84 kHz - 22.1 kHz)

Chen Yan's version: Between measures 10-1 to 17-1, the high-frequency components are clear and concentrated, with the blue-green area showing good extension, indicating excellent breath support and bright register transitions in the upper range. Particularly in the climactic section of the main melody of "Phoenix Tree," her head resonance control is stable, with a strong sense of reverberation.

Huang Lu's version: In the same section, the high-frequency distribution appears relatively sparse and less sustained, reflecting a more conservative approach in the upper range, often using mixed or covered resonance techniques, sacrificing some penetration power to enhance emotional softness.

Mid-frequency range (2.21 kHz–8.84 kHz)

Chen Yan: This frequency range shows consistently stable green distribution with clear timbral characteristics and high speech intelligibility. Particularly in the "emotional transition zone" between measures 5-1 to 13-1, the mid-frequency peaks remain stable, best demonstrating the coordination of articulation and resonance in her performance.

Huang Lu: The mid-frequency range appears fuller, with distinct green "concentration zones" notably between measures 7-1 to 14-1,

DOI: 10.5281/zenodo.17277096

indicating her strong ability in timbral consistency and resonance control. However, compared to Chen Yan, the frequency variations are slightly less pronounced, showing relatively steady dynamic layers.

Low-frequency range (0–2.21 kHz)

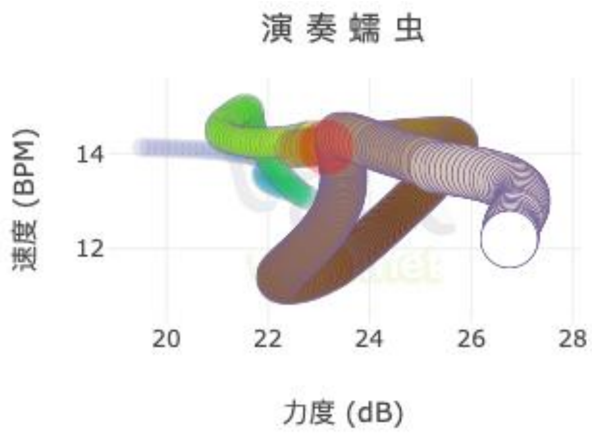
Chen Yan: The balanced distribution of red and orange areas indicates stable lower register usage with appropriate glottal closure. The low-frequency energy remains consistent throughout the piece without noticeable gaps, making it well-suited for emphasizing a "narrative" tone.

Huang Lu: The low-frequency performance is slightly weaker in measures 1-1 to 6-1 but strengthens noticeably in measures 10-1 to 15-1. This likely reflects a deliberate strategy to enhance emotional expression by employing chest resonance to convey a "whispered" sentiment.

Chen Yan's high-frequency treatment demonstrates greater penetration and spatial presence, making it ideal for lyrical and expressive passages, while Huang Lu adopts a more restrained and introspective approach with subtler emotional delivery. In the mid-frequency range, Chen Yan emphasizes clarity and dimensionality, whereas Huang Lu prioritizes softness and uniformity. In the low-frequency range, Chen Yan maintains balanced control, while Huang Lu strategically enhances specific sections to create intimacy, showcasing her stronger dramatic construction skills. Between measures 10-1 to 13-1, Chen Yan exhibits a broader frequency span (extending from 2kHz to 19kHz), reflecting her bold handling of register transitions through mixed head and chest voice techniques, creating an aesthetic tension of "timbre leaps." In contrast, Huang Lu opts for a narrower frequency band in the same passage, demonstrating finer transitional control and a focus on natural phrasing and linguistic flow. From the perspective of spectrogram analysis, the two vocalists Chen Yan and Huang Lu represent two distinct interpretive approaches to "Phoenix Tree": one emphasizes both technical mastery and dramatic tension while focusing on structural variations (Chen Yan), while the other features introspective emotional flow with refined articulation (Huang Lu). Through detailed analysis of frequency, energy, resonance and other dimensions, we can gain deeper insights into the personalized performance approaches in Chinese art songs, while also providing both technical and aesthetic references for teaching and performance.

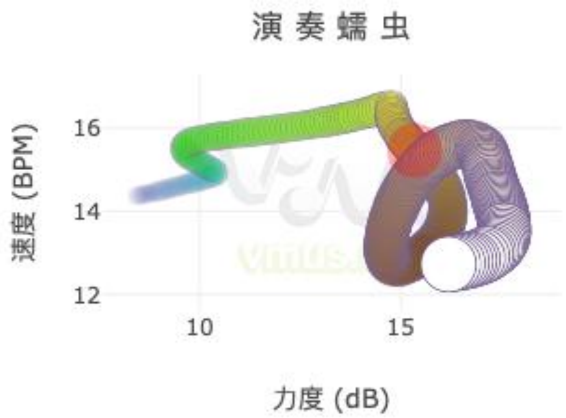
The Music Worm Plot is a three-dimensional visualization method that integrates time, tempo, and dynamics, where the X-axis represents dynamics (dB) indicating volume intensity, the Y-axis represents tempo (BPM) reflecting the speed of performance, and the color variations and path trajectories reveal the evolving emotional tension and dynamic control during the vocal performance. This visualization effectively captures nuanced aspects such as breath control, rhythmic flexibility, and emotional fluctuations, which are otherwise challenging to quantify, making it particularly valuable for comparative analysis.

Figure*10. Vmus.net CHEN YAN'Phoenix Tree'Music Worm Plot*



Source: *Vmus.net*

Figure11. *Vmus.net HUANG LU'Phoenix Tree'Music Worm Plot*



Source: *Vmus.net*

In terms of dynamic control (X-axis), Chen Yan's dynamics range predominantly between 22dB-28dB, with higher average levels indicating robust breath support and resonance techniques (likely blending chest and head resonance) to emphasize emotional depth. Key phrases such as"我站在梧桐树下"(I stand beneath the phoenxi tree") exhibit intensified dynamics for dramatic expression. For Huang Lu, HER dynamic range is concentrated between 9dB to 16dB, generally softer, demonstrating more delicate and gentle timbre control, presenting a more "introverted" singing style that emphasizes refined phrasing and articulation. The two singers employ two distinct interpretive approaches: Chen Yan exhibits a "broad dynamic range" expressive strategy, while Huang Lu adopts a "local micro-dynamic treatment" as the core of her singing technique.

Regarding tempo control (Y-axis direction), Chen Yan's tempo fluctuates approximately between 12.5 and 15.5 BPM, with a trajectory showing three downward-upward-downward shifts, indicating her frequent use of rubato, particularly slowing down at emotional climaxes ("dragging" the beat) to enhance expressiveness. Her approach demonstrates greater rhythmic flexibility. Huang Lu's tempo remains stable between 13 and 16 BPM with minimal fluctuations, reflecting a more metrically anchored performance that prioritizes structural clarity and rhythmic continuity. Her style leans toward rhythmic control.

As for path structure and musical tension, Chen Yan's worm plot features thick, spiraling trajectories with three-phase tension buildup, demonstrating a layered expressive approach - a "progressive-tension vocal architecture" that emphasizes structural intensification like poetic progression. Huang Lu's worm plot shows

DOI: 10.5281/zenodo.17277096

a diagonally ascending path with minor fluctuations, indicating a "stable-gradual emotional delivery" where tension is achieved in one cohesive motion, focusing on holistic emotional development akin to ink-wash painting's fluid subtlety.

Analysis Dimension	Chen Yan	Huang Lu
Dynamic Range	22~28 dB, wide and expressive	9~16 dB, concentrated and delicately soft
Tempo Fluctuation	Significant variation, free rhythm	Minimal variation, steady rhythm
Performance Style	Emotionally intense with strong breath support	Restrained and introspective with smooth phrasing
Resonance Strategy	Mixed resonance for enhanced tension	Bright head-dominant resonance pursuing clarity
Tension Construction	Multi-stage progressive development	Linear flowing progression

Quantitative analysis through Vmus.net's Music Worm Plot clearly reveals the underlying dynamic tension and rhythmic structure in vocal performances, which holds significant importance for comparing artistic styles and individual interpretation approaches in art songs.

Conclusion

This study focuses on the performance styles of the Chinese art song "Phoenix Tree," employing music visualization techniques to explore new approaches for vocal performance analysis. The study begins by categorizing current music visualization technologies, clarifying their application value in music structure analysis and performance style comparison, with particular emphasis on the functional features and operational advantages of Vmus.net software. Subsequently, using two performance versions of "Phoenix Tree" by Chen Yan and Huang Lu as research subjects, the study conducts a comparative analysis across multiple dimensions—including rhythmic control, emotional expression, frequency energy distribution, and resonance strategies—by integrating tempo-dynamics plots and Spectrogram. The results indicate that Chen Yan's version is more stable and refined, with concentrated frequency spectra, reflecting the discipline of traditional academic training, while Huang Lu's one exhibits greater variation in tempo and dynamics, with broader frequency distribution, demonstrating distinct personalized expression. The research demonstrates that music visualization not only helps reveal performers' technical and expressive characteristics in depth but also provides scientific and visual support for the teaching and analysis of art song performance.

REFERENCES

1. MA LING. 情感表达在声乐演唱中的重要性及其表达手法探究 [J]. 四川戏剧 ,2015,(04):127-129
2. GAO LIN . 浅谈歌曲《梧桐树》演唱过程中的情感表达 . XIA MEN University
3. WANG ZI YAO. 《梧桐树》歌曲演唱中的情感表达 [J].
4. YANG JIAN. www.Vmus.net .V1.1 manual and V2.0 introduction
5. YANG JIAN、HUANG YING .可视化分析在音乐表演研究中的综合运用.《中央音乐学院学报》2016年第4期46-58页
6. HUANG YING、YANG JIAN .音乐表演中的姿态与协作——基于音视频媒体的可视化分析.《黄钟(武汉音乐学院学报)》2022年第1期73-83页
7. A. R. Jensenius, M. M. Wanderley:"Musical Gestures: Concepts and Methods in Research", edited by R. I. Godøy and M. Leman: Musical Gestures: Sound, Movement, and Meaning, London:Routledge, 2010, pp.24-47. 赖菁菁：《表演姿态的概念与分类——从跨学科的音乐姿态研究视角出发》，《人民音乐》2014年第4期，第46—49页。
8. J. W. Davidson: "Bodily Movement and Facial Actions in ExpressiveMusical Performance by Solo and Duo Instrumentalists: Two Distinctive Case Studies", Psychology of Music, 2012,40(5), pp. 595-633.
9. J.W. Davidson: "Bodily Movement and Facial Actions in ExpressiveMusical Performance by Solo and Duo Instrumentalists: TwoDistinctive Case Studies", Psychology of Music, 2012,40(5), p. 616.

DOI: 10.5281/zenodo.17277096