



# 動画をつくるための インタフェースデザイン

## User Interface Design of Video Authoring Tools

産業技術総合研究所／アーチ 加藤 淳

### 1. はじめに

動画と一口に言っても、さまざまな様式があり、用途があり、楽しみ方がある。映画館で見る実写映画、電車内のディスプレイに映る広告、テレビで流

れるニュース、VR ヘッドセットで楽しむ全天周映像、スマートフォンで視聴するミュージックビデオ（音楽動画）や VTuber の配信、家族の成長記録に至るまで、私たちの日々の生活は動画で彩られてい

### 著作物利用に関する注意

本著作物の著作権は特定非営利活動法人ヒューマンインタフェース学会に帰属します。本著作物は著作権者であるヒューマンインタフェース学会の許可のもとに掲載するものです。ご利用に当たっては「著作権法」に従うことをお願い致します。

### 引用情報 (<https://junkato.jp/ja/publications>)

加藤 淳 (2023). 動画をつくるためのインタフェースデザイン. ヒューマンインタフェース学会誌, 25(2), 12-15.

る。そして、動画制作も、もはや一握りのプロが高価な機材で取り組むものではなくてきている。その要因として、本稿では、ソフトウェアの動画制作ツールが使いやすい制作支援インタフェースを備えるようになってきたことに注目したい。とくに、著者がこれまで取り組んできた音楽動画やアニメの制作支援研究の知見も織り交ぜながら、快適な動画制作を支えるインタフェースデザインの一端を紹介する。そして、インタフェースデザインがクリエイタ同士の共創を後押しすることの重要性について述べる。

## 2. 誰でも動画をつくれるインタフェースデザイン

スマートフォンやタブレットなどの標準機能（例えば iPhone の「写真」アプリ）では、動画の再生画面に「編集」ボタンが表示されており、スムーズに編集画面に移行できる。こうした一般向けのインタフェースデザインでは、動画の再生から編集と共有までをストレスなく行えるようにすることが重視されている。複雑な機能は、ツールアイコンのタップなど明示的な操作時のみ表示したり、あえて省いたりする工夫がなされている。

著者が研究開発してきた Web ベースの音楽動画制作支援サービス「TextAlive」(<https://textalive.jp>)<sup>[1]</sup>でも、楽曲を選択するだけで自

動生成された音楽動画を再生できる（図 1a）。演出の「スタイル」を切り替えられるパネルをタップすると、初めて、フォントの種類や配色などを調整できるウィジェットが表示される（トップページの画像）。さらに微調整がしたい場合だけ、典型的な動画編集用ツールのような画面へ遷移できる（図 1b）。このように多様なスキルセットを持った人々が利用するソフトウェアでは、単一のユーザインタフェース設計でなく、習熟度や実現したいタスクに応じて異なる詳細度のインタフェース（レイヤードインタフェース）を用意することが望ましいだろう。動画は視覚と聴覚を同時に刺激するマルチメディアであり、映像と音の同期が取れている必要がある。例えば、撮影した写真など、映像中で見せたいコンテンツが決まっている場合、動画の尺やコンテンツの切り替えペースが定まり、これを制約条件として、適切な音楽を選択したり自動生成したりといったインタラクションが考えられる。

著者らの TextAlive は音楽動画を制作するサービスであり、音として聴かせたいコンテンツが先に決まっている。そこで、音楽理解技術によって歌詞の発声タイミングやビート、サビの時刻情報などを自動的に抽出し、タイミングよく動くアニメーションを生成している。また、TextAlive や、TextAlive の連携サービスである Songle (<https://songle.jp>) では、さまざまな時刻情報やコード進行などを不特定多数で協力して、クラウドソーシングで訂正・管理するためのインタフェースを用意している（図 2）。一般に、コンテンツの自動理解技術は完璧ではない。推定結果が誤りを含んでいる場合に備えて訂正インタフェースを提供することは重要である。

動画は、最終的には特定の解像度とフレームレートの映像と特定のチャンネル数とサンプリングレートの音の組として書き出されるが、制作工程では多くの時系列データを組み合わせて演出を作り込んでいく。例えば TextAlive では、歌詞のテキストがフレーズ、単語、文字という階層構造で扱われており、映像を一コマ単位で編集するのではなく、フレーズごとに文字の動き方の「テンプレート」を選んだ



図 1 a) TextAlive で楽曲を選んだ直後の画面  
b) 詳細な動画編集用インタフェース



図 2 TextAlive の歌詞発声タイミング管理画面



り、単語ごとにフォントを変更したりできる（図 1b）。また、Adobe 社のベータ版製品 Adobe Blink（<https://labs.adobe.com/projects/blink>）では、ナレーションのテキストを編集すると、映像と音の対応する箇所も編集されるインタラクションが実現されている。今後、人工知能（AI）技術を活用した創造性支援ツールは、さらに人にとっての「意味」単位での、セマンティックな編集を可能にするインタフェースが充実する方向に進化していこう。

人にとって適切な意味単位や、そのためのインタフェースデザインを知りたい場合、コンピュータ導入前からある様式が参考になる。著者らは商業アニメ制作のプリプロダクション工程でデファクトスタンダードとして利用され、アニメの設計図とも呼ばれる「絵コンテ」の様式に着目し、文献調査とアニメ監督との緊密な連携により Web ベースの絵コンテ制作支援ツール「Griffith」<sup>[2]</sup>を設計、開発してきた（図 3）。

絵コンテは、アニメの監督・演出家が、多数のアニメーターなどが関与するプロダクション工程への指示内容をビジュアルにまとめたものである。絵のコマの隣に、誰がどこでどのような芝居をして何を話しているかといった文字情報と、「7+12」（7 秒+12 コマ＝7.5 秒）のようなタイミング情報とが書かれており、まさに、動画の内容をセマンティックかつコンパクトに表現する仕様書である。

多くの動画編集ツールでは、再生用シークバーが横方向であることの類推か、時間軸が横方向に伸びている。しかし、紙の絵コンテはコマが 4 コマ漫

画のように上から順に並ぶ。Griffith でも同様に時間軸を縦にしたところ、映像の左右バランスを確認しやすく、前後コマが手で隠れないなどのメリットを確認できた。

なお、紙の絵コンテは 1 枚あたり 4-6 コマしか書けないが、デジタルではそうした区切りは不要で、絵コンテ全体を縦に無限にスクロールできる（図 3a）。また、紙だと拡張はできないが、デジタルなら絵コンテ全体を一望するオーバービューを用意することもできる（図 3b）。アナログのツールをデジタル化する際は、ツールの使い勝手を単に再現するのではなく、ツールが表すクリエイターのセマンティクスをデジタルの文脈で解釈し直すことで、使いやすいインタフェースを実現できる。

### 3. みんなでつくる動画制作の未来

ここまで、TextAlive と Griffith という設計方針も用途も全く異なる事例を通して、動画をつくるためのインタフェースデザインのヒントを具体的かつ多角的に紹介してきた。最後に、著者がこれらのツールを設計する上で、共通して重視してきた点、すなわち共創を後押しするインタフェースデザインについて述べる。

音楽動画の制作工程には、作曲者、作詞者、動画制作者、イラストレーターなど多様な人々が携わっている。さらに TextAlive の場合、動画演出の「テンプレート」のアルゴリズムを開発できるインタフェースが用意されていて、プログラミングでツールの表現力を拡張できる（図 4a）。そこで TextAlive には、プログラムの実装に貢献した人々を含めて動画のクレジットを確認できるインタフェース（図 4b）があり、動画の書き出し時にもテキストファイルに明記される。アニメの制作工程も言うまでもなく共創であり、Griffith には、監督・演出家のための絵コンテ制作インタフェース

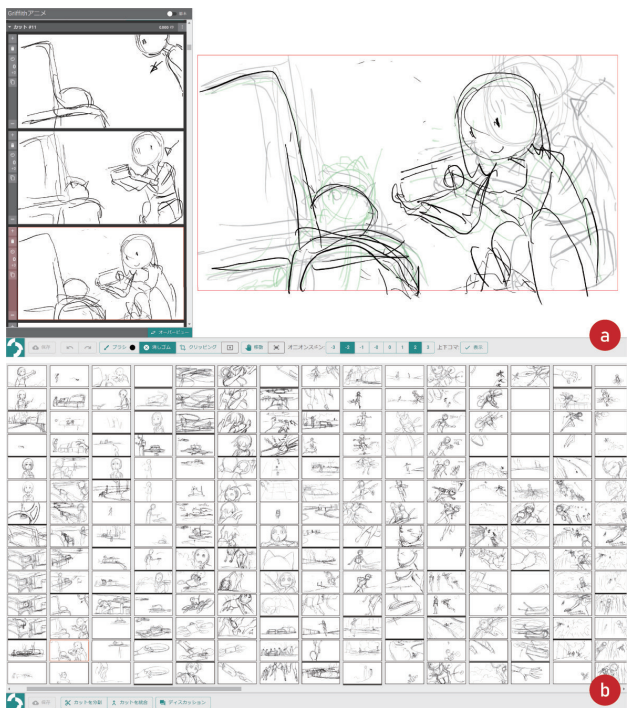


図 3 a) Griffith の絵コンテ編集画面  
b) 全体を一望できるオーバービューインタフェース

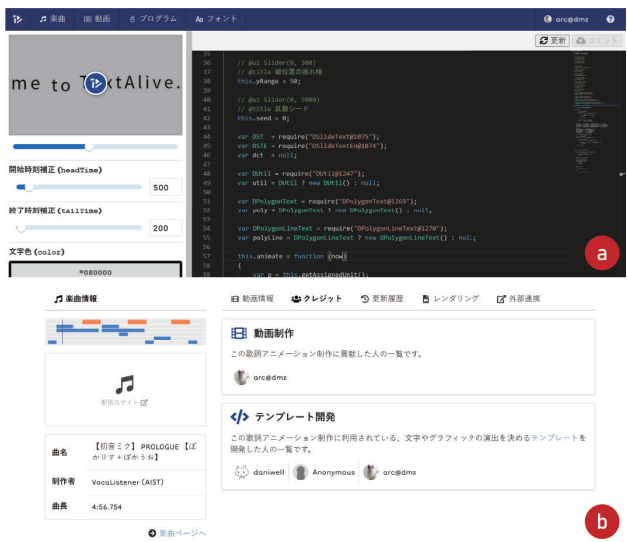


図 4 a) TextAlive のプログラミング用画面  
b) クレジット表記

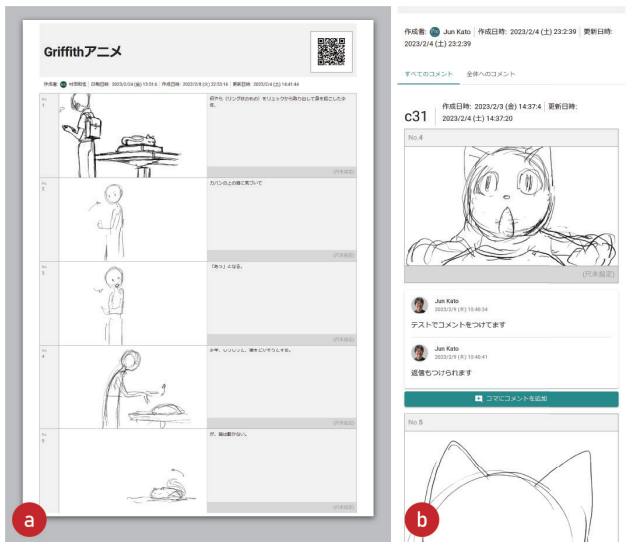


図 5 a) Griffith の印刷用画面  
b) スマートフォン用の閲覧・議論インタフェース

の他に、絵コンテのデータに対し多様なインタラクションが可能なインタフェースが独立して用意されている。例えば、プリプロダクション工程の幅広い職種の人が使えるプリビズ（絵コンテをパラパラ漫画風に再生する機能）や、プロダクション工程のアニメーターなどが紙に印刷するためのインタフェース（図 5a）やスマートフォンで閲覧したりコメントを残したりできるインタフェース（図 5b）がある。

TextAlive の動画編集とプログラミングのインタフェースや Griffith の職能別のインタフェースは、動画に関する情報を、スキルセットや権限に応じて適切に利活用できるようにするものである。また、コメントを残せるようにしたりクレジットを明示したりする機能は、共創における直接間接のコミュニケーションを後押しする。このように、動画制作を楽しく創造的でフェアな共創体験にする上で、ツールの機能設計とインタフェースデザインが担う役割は非常に大きい。

さらに広い視野に立てば、筆者らのツール開発自体、ユーザの創作に後押しされている。2015 年の一般公開以来、TextAlive で制作された動画を共有サービスで見かけることが増えており、想



a Summer lights / 市瀬るぼ feat. 鏡音リン  
<https://youtu.be/PT0pqQsGFaM>



b ORI01 feat. Hatsune Miku / BIGHEAD  
<https://youtu.be/-f1lRZuYf>



c Nagaoka Mikunological Live 2022 in 技大祭  
<https://youtu.be/OiV0nExib9g>

図 6 TextAlive で制作された音楽動画の様々な応用例

像を超える応用を何度も目の当たりにして刺激を受けてきた。例えば動画を加工してイラストと効果的に組み合わせた作例（図 6a）や、仮想現実（VR）のコンテンツ制作ツールで動画を読み込んで 3DCG 中に配置した作例（図 6b）、バーチャルキャラクターの 3DCG ライブで紗幕に投影した例（図 6c）などである。動画をはじめとするメディアの価値は、新たな表現手段と組み合わせることで、何度でも生まれ変わる。これからも、クリエイターの方々と手を携えながら、技術の力でメディア表現の未来を切り拓いていきたい。

加藤 淳 かとう じゅん

博士（情報理工学）。2014 年より国立研究開発法人 産業技術総合研究所研究員、2018 年より同主任研究員。同年よりアーチ株式会社技術顧問を兼務。 Human-Computer Interaction、とくにプログラミング体験向上など人々の創造的活動を支援する研究に従事し、クリエイター・プログラマー向け創造性支援ツール・サービスを開発、運営。詳細は <https://junkato.jp/ja> 参照。



### 参考文献

- [1] Jun Kato, Tomoyasu Nakano, Masataka Goto. TextAlive: Integrated Design Environment for Kinetic Typography. In Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI 2015), pp.3403-3412, 2015. <https://junkato.jp/ja/textalive>
- [2] Jun Kato, Ryotaro Mihara, Nao Hirasawa. Past, Present, and Future of Storyboarding in Japanese Animation. In Society for Animation Studies 2021 (SAS 2021), 2021. <https://research.archinc.jp/griffith>