

アプリ仕様書（ドラフト）

製品名: Vocal Mastery Lab (VML)

ネーミング検討（最終候補）

Vocal Mastery Lab / Vocal Precision Lab / Vocal Refinery Lab ここに比較・推奨を追記しました（§ 0）。

本アプリは、ユーザーが合法的に所持する音源からボーカルトラックを抽出し、プロ歌手の歌唱と自分の歌唱を比較分析しながら学習できる歌唱トレーニングアプリである。以下に詳細仕様を整理する。

1. アプリ概要

- ・ **アプリ目的:** プロの歌い方を科学的に分解し、ユーザーの歌唱を比較分析して改善につなげる
- ・ **主要特徴:**
 - ・ 音源からボーカル抽出（オンデバイス処理）
 - ・ プロ歌唱の歌い方解析（F0/ビブラート/タイミング/声質）
 - ・ ユーザー歌唱との比較・フィードバック
 - ・ カラオケ練習モード + 分析モード

2. 想定ユーザー

- ・ 歌唱を改善したい一般ユーザー
- ・ ボーカルトレーニング用途
- ・ 楽曲研究者・音楽教育分野

3. ユースケース

- ・ プロ歌手の歌唱を参考に自分の歌い方を調整したい
- ・ カラオケ採点では得られない精密な歌唱分析を行いたい
- ・ ピッチや表現技術（ビブラート、抑揚）を学びたい

4. 機能一覧（概要）

- ・ 楽曲インポート（DRMチェック含む）
 - ・ ボーカル分離（Vocals/伴奏）
 - ・ プロ歌唱解析・構造化
 - ・ ユーザー歌唱録音
 - ・ 比較分析（ピッチ・タイミング・ビブラート・声質）
 - ・ AIフィードバック
 - ・ 練習進捗管理
-

5. アプリ機能仕様（メイン）

5.1 対応プラットフォーム / 非機能要件

- 対象デバイス: iOS (iPhone・iPad)
- 動作方針: 完全オフライン (音源・特徴量・結果を外部送信しない)
- 精度優先: モデルは精度寄りの構成 (MDX/HTDemucs系の高精度モデルをCore ML化)
- 最小OS: iOS 16 以上 (Metal Performance Shaders / Core ML最適化前提)
- 推奨端末: A16 Bionic 以降 (iPhone 14 Pro~)
- 端末負荷: バックグラウンド処理と優先度制御 (処理中でもUIが滑らか)

5.2 ユーザーフロー（要約）

1) 音源選択 (Files / ミュージックライブラリ) → DRMチェック 2) 区間選択 (5~20秒) → プレビュー 3) 分離 (一度で Vocals と Accompaniment を得る) 4) リファレンス解析 (Vocals: F0/ビブラート/スペクトル) 5) 練習 (Accompaniment再生+ユーザー録音) 6) 比較解析 (リファレンス vs ユーザー) 7) フィードバック表示 (スコア/指標/コメント)

5.3 画面仕様（ドラフト）

- ホーム: 最近のセッション、読み込みボタン
- 区間エディタ: 波形・スペクトログラム、IN/OUTハンドル、再生ループ
- 分析ビュー: スペクトログラム+F0オーバーレイ、ビブラート帯可視化、ミキサー (Vocals/Accomp)
- 練習ビュー: 伴奏再生、録音、メトロノーム、進行バー、先読みインジケータ
- 結果ビュー: KPIカード (ピッチ・タイミング・ビブラート・声色安定度) + 差分ヒートマップ

5.4 音声処理パイプライン（精度優先）

- リサンプリング: 48kHz mono (解析系はmonoで統一)
- STFT設定: nFFT=2048, hop=256 (解析) / 分離モデルに準拠
- 分離モデル: MDX-Net (精度版)、必要に応じHTDemucsも検討
- チャンク処理: 10秒/窓、1秒オーバーラップ、ハンニング窓で合成
- Vocals解析:
 - F0: CREPE(Core ML) または AudioKit PitchTap→最終的にCREPEへ移行
 - Vibrato: F0(cent)のトレンド除去→4-8Hz BPF→レート/エクステント/安定度
 - スペクトル: セントロイド/ロールオフ/フラックス/MFCC(1-13)
 - アラインメント: DTW (コスト= $|F0_ref - F0_user| + \lambda \cdot |\Delta E|$) で時間整合

5.5 スコアリング (β)

フィードバック仕様は要検討のため、暫定の定義を用意。後日チューニング。

- **Pitch Accuracy**: 中央値・95%分位 (cents)
- **Timing Deviation**: ノート境界差(ms)
- **Vibrato**: Rate(Hz)・Extent(cents)・Stability(CV)
- **Timbre Consistency**: MFCC分散、セントロイド分散

総合スコア: 重み $w = \{0.4, 0.2, 0.2, 0.2\}$ を初期値 (可変)

5.6 データ構造（抜粋）

```
{
  "session_id": "ISO8601",
  "source": {"url": "file://...", "drm": false},
```

```

"segment": {"start": 42.5, "duration": 12.0},
"features": {
  "ref": {"f0_hz": [..], "vibrato": {"rate": 5.8, "extent": 70, "stability": 0.15}},
  "user": {"f0_hz": [..], "env_db": [..], "mfcc": [[..]]}
},
"scores": {"pitch_med_c": 22.4, "timing_ms": 40, "vibrato_rate": 5.9}
}

```

5.7 モジュール構成

- **SeparationEngine** (Core ML推論 + STFT/iSTFT)
- **FeatureExtractor** (F0/スペクトル/包絡)
- **Alignment** (DTW/NCC)
- **Scorer** (KPI算出)
- **SessionStore** (JSON保存・暗号化)
- **AudioIO** (再生・録音・AEC)

5.8 パフォーマンス目安 (精度構成)

- iPhone 15 Pro: 分離 **10秒** ≈ **1-2秒**、解析リアルタイム以下
- iPhone 13: 分離 **10秒** ≈ **2-4秒** (UIに先読みインジケータ)

5.9 安全設計

- ファイルはApp Sandbox内一時領域、セッション終了で自動削除 (設定で保存可否)
- バックグラウンドタスク中もデータ送信なし (Network権限不要)
- クラッシュ時もバッファは暗号化一時領域にのみ残存 (次回起動でパージ)

6. UI/UX 詳細仕様 (プロ志向 → 大衆向けの順に展開)

6.1 ナビゲーション構造

- **タブ構成 (下部)** : **ホーム** / **分析** / **練習** / **履歴** / **設定**
- **主要導線**: ホームの「新規セッション作成」 → 区間エディタ → 分離/分析 → 練習 → 結果
- **セッション概念**: 各セッションは「曲URL + 区間 + 解析結果 + 録音」を持つ

6.2 画面別仕様

A. ホーム

- コンテンツ: 最近のセッションカード、"新規セッション" FAB、チュートリアルリール
- 状態: Empty (未作成時にミニチュートリアル表示) / List / 検索
- 操作: ファイルインポート、ライブラリ選択 (DRMチェック即時)
- エラー: DRM検知時は**ブロックダイアログ** + ヘルプリンク

B. 区間エディタ (Wave/Spec)

- ビュー: 上段**波形**、下段**スペクトログラム** (48kHz / nFFT 1024, hop 256 の軽量表示)
- 操作: IN/OUT ハンドル、ズーム、ループ、スナップ (拍/小節)
- メタ: テンポ推定 (簡易BPM)、キー推定 (クロマ)

- UI要素：再生ボタン、ループ切替、BPM/キー表示、右上に"分離+分析" CTA
- ローディング：プログレスバー+キャプション（例："解析中：スペクトログラム生成"）

C. 分離 & 分析ビュー

- レイアウト：上**プログレス**（チャンクごとに進行）、中**Spec + F0(白)/振幅(灰)**オーバーレイ、下**ミキサー**
- ミキサー：**Vocals** / **Accompaniment** 独立スライダー（-∞〜+6dB）、モノラルチェック
- ステータス：**待機** / **分離中** / **分析中** / **完了**
- 失敗時：リトライ、軽量モデルへのフォールバック、区間短縮の提案

D. 練習ビュー

- プレイヤー：伴奏再生、拍子メトロノーム、カウントイン（1小節）
- 録音：AECオン（**voiceChat**）、入力レベルメーター、クリッピング警告
- ライブ表示：**ユーザーF0（青）**をリアルタイムでオーバーレイ（プロF0=白、不一致は赤ハイライト）
- 先読み：次チャンクをバックグラウンド分離、ラグを抑制
- ループ：AB ループ、速度可変（0.5x/0.75x/1.0x）

E. 結果ビュー（比較/フィードバック）

- 上段：Spec + **F0_ref(白)/F0_user(青)**、誤差ヒートマップ（±50c範囲を色分け）
- KPIカード：Pitch（中央値/95%）、Timing（ms）、Vibrato（Rate/Extent/Stability）、Timbre（MFCC安定度）
- コメント例：
 - 「1小節目頭拍で +32c → 先上がり傾向。次は頭拍を低めに意識」
 - 「ビブラートが 3.2Hz → もう少し速く（5-7Hz）に」
- シェア：**数値のみ**（音源・録音は共有不可）。PNGグラフにテキストを埋め込み

F. 履歴

- リスト：セッション名/日付/区間/主要スコア
- 詳細：結果ビューを再現、再練習へ導線
- 保存方針：デフォルトは**特微量とメタのみ**、音声は破棄（設定で保持ON可）

G. 設定

- オーディオ：入出力デバイス、レイテンシ、AEC ON/OFF
- 分離：精度/速度プリセット（**Pro 高精度** / **Balanced** / **Fast**）
- 解析：F0推定器（PitchTap/CREPE）、サンプルレート固定（48kHz）
- プライバシー：データ送信なしの明示、セッション自動削除期間
- 法務：利用規約/ライセンス/DRM説明

6.3 UX原則

- **予測可能**：処理は"チャンク進捗+残り目安"で可視化
- **中断に強い**：バックグラウンドで処理を継続、復帰でシームレス
- **誤作動耐性**：ノイズ・無声区間のマスキング、F0補間
- **アクセシビリティ**：ダークモード、VoiceOver、色覚多様性配慮（青/白/赤コントラスト）

6.4 エンpty/エラー/境界ケース

- ・DRM音源：即ブロック→ガイド（"購入済みDRMなし音源のみ対応"）
- ・分離失敗：区間短縮/軽量モデル切替のサジェスト
- ・マイク拒否：録音前チュートリアル+権限再リクエスト
- ・端末負荷：温度/バッテリー高時は自動で **Balanced** → フォールバック

6.5 計測と品質ゲート（オンデバイス）

- ・指標：分離1チャンクのレイテンシ、F0落ち率、無声検出精度
- ・ログ：匿名・デバイス内のみ（外部送信なし）
- ・クラッシュ復旧：未完了セッションを安全に巻き戻し

7. プロ志向 → 大衆向けの展開計画（今回リリースは“Pro志向版”）

7.0 位置づけの明確化（対比表）

項目	Pro志向（今回）	大衆向け（将来）
ターゲット	ボーカルトレーナー/上級者/研究寄り	初心者/ライト層/カラオケユーザー
体験重心	分析の正確性・再現性	速さ・分かりやすさ・楽しさ
モデル	高精度（MDX/HTDemucs精度版, CREPE）	軽量（mobile向け）
分析指標	F0/Timing/Vibrato/Timbre/MFCC	Pitch/Timingの要約＋一言アドバイス
画面	波形/Spec/F0重ね・KPI多め	色とゲージ中心・説明最小
所要時間	区間10秒の分離1-2秒（iPhone 15 Pro）	即時プレビュー優先（必要時のみ詳細処理）
価格	買い切り/Pro機能課金	フリーミアム/サブスク
サポート	詳細ヘルプ/ツールチップ/エキスパート設定	チュートリアル/クエスト/バッジ

7.1 フェーズ1（今回リリース：Pro向け・本格仕様）

- ・位置づけ：プロ/上級ユーザーの歌唱研究と精密トレーニングに特化
- ・核となる価値：原曲ボーカルの精密解析とユーザー歌唱の定量比較（F0/タイミング/ビブラート/声色）
- ・トーン&メッセージ：
- ・Storeコピー例：
 - 「プロの歌い方を“数値”で学ぶ。分離・解析・比較まで、すべて端末内で。」
 - 「カラオケ採点ではわからない“歌い方”を解剖して身につける。」
- ・端末要件：A16以降推奨/精度モデル標準
- ・機能：詳細分析（ビブラート/スペクトル）、高度ミキサー、区間細分、DTW比較、スコア詳細
- ・UI：情報量多め/専門指標の説明ツールチップ/ショートカット多数
- ・価格：買い切り＋Pro拡張（候補）

7.2 フェーズ2（将来リリース予定：大衆向け・簡易仕様）

- 位置づけ：誰でも1タップで“うまく歌える”体験（学習曲線を最小化）
- コア価値：結果の“意味”を直感表示（ライトなスコア＋一言アドバイス）
- トーン&メッセージ：
- Storeコピー例：
 - 「好きな曲で1分トレーニング。聴いて、真似して、すぐ上達。」
 - 「難しい言葉は不要。色とゲージで“今より良い”がわかる。」
- 端末要件：A15以降想定／軽量モデル標準
- 機能：ワンタップ分離・クイック分析・ガイド音・ゲーム化要素
- UI：シンプルUI／指標は数個に限定／説明は最低限

7.3 マーケ/ASOの文言ガイド（差別化が一目で伝わる）

- キーワード：
- Pro版："On-device", "Vocal Separation", "F0 / Vibrato", "DTW Alignment", "Research-grade"
- 大衆版："1-Tap", "Quick Practice", "Easy Score", "Sing Better"
- スクリーンショット方針：
- Pro版：Spec+F0の重ね表示、KPIカード、Vocals/Accompミキサー
- 大衆版：大きなゲージと色、ワンタップ操作、短い説明文

7.4 機能フラグ（同一コードで両対応）

- `FeatureFlags.proAnalysis`（詳細KPI/DTW）
 - `FeatureFlags.lightMode`（UI簡易＋軽量モデル）
 - `FeatureFlags.quickPreview`（M/Sの即時プレビュー→高品質レンダリングへ）
-

8. 権利対応ポリシー（法務・審査対応）

（以下、既存の権利セクションに続く）（法務・審査対応）（以下、既存の権利セクションに続く）