Gabor dual windows using convex optimization

Nathanaël Perraudin, Nicki Holighaus, Peter L. Søndergaard and Peter Balazs
Acoustics Research Institute, Austrian Academy of Sciences, Wohllebengasse 12–14, 1040 Vienna, Austria
Email: nathanael.perraudin@epfl.ch, nicki.holighaus@oeaw.ac.at, soender@kfs.oeaw.ac.at, peter.balazs@oeaw.ac.at

Abstract—Redundant Gabor frames admit an infinite number of dual frames, yet only the canonical dual Gabor system, constructed from the minimal $\ell^2$-norm dual window, is widely used. This window function however, might lack desirable properties, such as good time-frequency concentration, small support or smoothness. We employ convex optimization methods to design dual windows satisfying the Wexler-Raz equations and optimizing various constraints. Numerical experiments show that alternate dual windows with considerably improved features can be found.

I. INTRODUCTION

Time-frequency representations, in particular Gabor transforms [9], i.e. sampled Short-Time Fourier transforms, are ubiquitous in signal processing. Gabor transforms represent a signal as linear combination of translates and modulations of a single window function, which for best results should be chosen to be well-concentrated in time and frequency.

A signal can be reconstructed from its Gabor transform using a system with the same modulation and translation structure. Moreover, infinitely many such systems exist if the Gabor transform is redundant. Finding a dual system with desirable properties given a prescribed analysis window is the topic of this paper.

More explicitly, for $g \in \ell^2(\mathbb{Z})$, and $a, M \in \mathbb{Z}$, we define the Gabor system

$$G(g, a, M) := \left\{ g_{m,n} = g[\cdot - na]e^{2\pi im/M} \right\}_{n \in \mathbb{Z}, m = 0, \ldots, M-1}.$$  \hfill (1)

If $G$ is also a frame [5], we refer to the system as a Gabor frame. For $f \in \ell^2(\mathbb{Z})$, the corresponding Gabor transform is given by

$$(Gf)[m + nM] = \langle f, g_{m,n} \rangle = \sum_{l \in \mathbb{Z}} f[l] \overline{g_{m,n}[l]},$$  \hfill (2)

with the analysis operator $G$ as given by the infinite matrix

$$G[m + nM, l] := G_{g,a,M}[m + nM, l] := g_{m,n}[l].$$

Gabor synthesis is performed by applying the adjoint of $G$ to a coefficient sequence $c \in \ell^2(\mathbb{Z})$. The action of the synthesis operator can be equivalently described as

$$f_{\text{syn}}[l] = (G^*c)[l] = \sum_{m,n} c[m + nM] g[l - na]e^{2\pi iml/M}. $$  \hfill (3)

The concatenation $S = G^*G$ of the analysis and synthesis operators is called the frame operator.

Reconstruction can be realized using the so-called canonical dual system, obtained by inverting $S$ and defined as

$$\tilde{g}_{m,n} = S^{-1}g_{m,n}. $$  \hfill (4)

In the particular case of Gabor frames, the canonical dual system is again a Gabor frame, i.e. it equals $G(\tilde{g}_{0,0}, a, M)$. Therefore we refer to $\tilde{g} = \tilde{g}_{0,0} = S^{-1}g$ as the canonical dual window.

The synthesis operator of $\tilde{g}$ coincides with the pseudo-inverse of the original analysis operator, i.e. $G^*_{\tilde{g},a,M} = G^\dagger$. So the inversion formula reads

$$f[l] = \sum_{m,n}(f,g_{m,n})\tilde{g}_{m,n}[l] = G^\dagger Gf[l]. $$  \hfill (5)

There are several approaches for finding the canonical dual in an efficient way, e.g. [4], [11]. Only if the length of the window $L_g$ is less than or equal to the number of channels $M$, is the canonical dual guaranteed to have the same length. This so-called painless case construction is omnipresent in signal processing, to the point where $M$ and $L_g$ are not distinguished.

Redundant Gabor frames possess infinitely many dual Gabor frames of the form $G(h, a, M)$, any of which facilitates perfect reconstruction from unmodified coefficients. On the other hand, whenever the coefficient representation is processed, varying dual systems provide different reconstructions and the features of the chosen system suddenly play an important role. Some of the ‘alternate duals’ might possess properties preferable to those of the canonical dual, e.g. shorter support, better localization or smoothness.

For a Gabor frame $G(h, a, M)$, the Wexler-Raz equations [17], [20] provide a necessary and sufficient condition to constitute a dual frame for $G(g, a, M)$. Using this hard constraint, a convex optimization problem can be defined by adding functionals to be minimized that provide desired properties.

Recently, convex optimization in the context of audio signal processing has grown into a active field of research and in particular proximal splitting methods [6], [7], [8] have been used to great effect, e.g. in audio inpainting [2], [1] and sparse representation [12]. In those cases, optimization techniques are applied directly to the signal or its time-frequency representation. In this contribution, we apply optimization techniques to shape the building blocks of the time-frequency representation instead. Since a systematic evaluation of the available optimization techniques is beyond the scope of this contribution, we only present an exemplary realization.

Our method is a much more general approach than the construction of non-canonical dual windows found in [19] and optimizes several criteria at once. One particular application of the proposed approach is the construction of smooth dual windows satisfying a support constraint. To illustrate the viability of our method, we choose a Gabor frame $G(g, a, M)$
with $g$ being an FIR window, i.e. a window function supported on a finite interval $I_g$, and construct a smooth dual window $h$ supported on an interval $I_h$.

II. GABOR FRAMES

In this contribution, we consider Gabor systems $G(g,a,M)$ in $\ell^2(\mathbb{Z})$. Such a system constitutes a frame if constants $0 < A < B < \infty$ exist, such that

$$A\|f\|_2^2 \leq \|Gf\|_2^2 \leq B\|f\|_2^2,$$

for all $f \in \ell^2(\mathbb{Z})$. (6)

In that case, the closed linear span of its elements equals $\ell^2(\mathbb{Z})$ and every sequence $f \in \ell^2(\mathbb{Z})$ can be written as

$$f = G^*c,$$

(7)

for some coefficient sequence $c \in \ell^2(\mathbb{Z})$. In particular, if $G(h,a,M)$ is a dual Gabor frame, $c = G_{h,a,M}f$ is one possible choice. Note that frames are “mutually dual”, i.e. the role of $G(g,a,M)$ and $G(h,a,M)$ in the considerations above can be switched at will.

The Wexler-Raz equations [20], [17] for $\ell^2(\mathbb{Z})$ provide a necessary and sufficient condition for a function $h \in \ell^2(\mathbb{Z})$ to be a dual Gabor window for $G(g,a,M)$. They are given by

$$M \sum_{a} \delta(h - nM) = \delta[n]h[n],$$

(8)

for $m = 0, \ldots, a-1, n \in \mathbb{Z}$. In the equation above, $\delta[l]$ denotes the Kronecker delta at position $l$. In terms of the analysis matrix $G^o = G_{g,M,a}$, i.e. switching the role of $a$ and $M$, they can be stated as

$$G^o h = \frac{a}{M} \delta.$$

(9)

III. PROXIMAL SPLITTING METHODS

The convex optimization problems we consider are of the form

$$\min_{x \in \mathbb{R}^L} \sum_{i=1}^{K} f_i(x),$$

(10)

where the $f_i$ are convex functions. Note that if at least one function $f_i$ is not differentiable, it is not possible to apply smooth optimization techniques. Proximal splitting methods [7], on the other hand may still apply. The term proximal splitting originates from the fact that each function $f_i$ is minimized iteratively with the help of their corresponding proximity operator, a generalization of convex projection operators, defined as follows.

Definition 1. The proximity operator of a function $f \in \Gamma_0(\mathbb{R}^L)$ is defined by

$$\text{prox}_f(y) := \arg\min_{x \in \mathbb{R}^L} \left\{ \frac{1}{2}\|y - x\|^2 + f(x) \right\}.$$

(11)

Since $f$ is convex, the minimization problem in (11) has a unique solution for every $y \in \mathbb{R}^L$ and consequently $\text{prox}_f : \mathbb{R}^L \rightarrow \mathbb{R}^L$ is well-defined.

More information on the properties of proximity operators can be found in [16], [13].

From now on, we will denote by $i_C$ the indicator function [7], of a non-empty, closed and convex set $C \subset \mathbb{R}^L$ by

$$i_C : \mathbb{R}^L \rightarrow \{0, +\infty\} : x \mapsto \begin{cases} 0, & \text{if } x \in C \\ +\infty, & \text{otherwise} \end{cases}$$

(12)

and by $\Gamma_0(\mathbb{R}^L)$ the class of functions

$$\Gamma_0(\mathbb{R}^L) = \{ f : \mathbb{R}^L \rightarrow \mathbb{R} : f \text{ lower semi-continuous, convex and proper} \}.$$

Indicator functions can be used to add hard constraints, e.g. a set of linear equations that the solution must satisfy, to an optimization problem of the form (10). More explicitly,

$$\arg\min_{x \in \mathbb{R}^L} \sum_{i=1}^{K} \lambda_i f_i(x) = \arg\min_{x \in \mathbb{R}^L} \sum_{i=1}^{K} \lambda_i f_i(x) + i_C,$$

(13)

where $C = \{ x \in \mathbb{R}^L : x \text{ satisfies the hard constraints} \}$ is the set of admissible points. If $C$ is non-empty and convex, Equation (13) has a solution for any given choice of regularization parameters $\lambda_i$, uniquely determined if at least one $f_i$ is strictly convex.

Table I presents a list of commonly used regularizer functions $f_i$ that can be combined to tune the solution $x$.

<table>
<thead>
<tr>
<th>Function</th>
<th>Effect on the signal</th>
</tr>
</thead>
<tbody>
<tr>
<td>$|x|_1$</td>
<td>sparse representation in time</td>
</tr>
<tr>
<td>$|\mathcal{F}x|_1$</td>
<td>sparse representation in frequency</td>
</tr>
<tr>
<td>$|\nabla x|_2^2$</td>
<td>smooth representation in time / concentrated in frequency</td>
</tr>
<tr>
<td>$|\nabla \mathcal{F}x|_2^2$</td>
<td>smooth representation in frequency / concentrated in time</td>
</tr>
<tr>
<td>$|x|_2^2$</td>
<td>spread values more evenly</td>
</tr>
<tr>
<td>$\lambda(x)$</td>
<td>force $x \in C$</td>
</tr>
</tbody>
</table>

We decided to present a solution of (10) using the parallel proximal algorithm (PPXA, Algorithm 1). However, this contribution does not intend to propose the best method to solve (10), and other algorithms, e.g. generalized forward backward [15], might prove more efficient. Instead, we focus on a new formulation of the problem of finding dual Gabor windows.

In the next section we present one of the possible ways to solve (10). Optimality studies are beyond the scope of this paper and planned as future work.

IV. METHODS

Utilizing the theory established in the previous sections, we can now describe our method in detail. We intend to compute non-canonical dual windows for a given Gabor frame $G(g,a,M)$, where $g$ is an analysis windows supported on some finite interval $I_g$. Furthermore, we want the dual window to
be supported on an interval \( I_h \) and denote the convex set of all signals satisfying this constraint by \( C_{\text{supp}} \).

Considering the support constraint, we see that all but a small subset of the Wexler-Raz equations are trivially satisfied. Without loss of generality we assume \( I \) to be centered around 0. Noting that \( I \cap (I_h + nM) = \emptyset \) for \( |n| \geq \frac{L_g + L_h}{2M} \), only the equations for

\[
|n| < \frac{L_g + L_h}{2M}, \quad (14)
\]
can possibly be non-zero. This makes a total of \( 2a(\left\lfloor \frac{L_g + L_h}{2M} \right\rfloor + 1) \) equations in \( L_g \) unknowns. As a consequence, we are not required to consider sequences of infinite length to compute the dual window, but we can equivalently work with signals in \( \mathbb{C}^L \), where \( L \) is some multiple of \( a \) and \( M \) satisfying \( L \geq L_g + L_h + 1 \).

The solutions of the non-trivial equations from the Wexler-Raz equation system (8), numbered as in (14) form a convex set \( C_{\text{dual}} \), providing the second hard constraint after the support condition.

Then, \( C = C_{\text{dual}} \cap C_{\text{supp}} \) is also convex and if non-empty,\(^1\) forms a legal set of admissible points for a problem of the form (13). To shape the resulting dual window towards some useful properties, we select suitable regularization functions (Table I) and parameters, employing PPXA to solve the resulting convex optimization problem, converging to the unique solution. The indicator functions \( i_{C_{\text{dual}}} \) and \( i_{C_{\text{supp}}} \) are used to realize the duality and support constraints.

Experience shows that PPXA needs a large number of iterations to perfectly satisfy the hard constraints. To speed up this process, a final projection is performed once the algorithm converges to a certain accuracy. If there is more than one regularization function to be minimized, the projection is realized by a POCS (Projection Onto Convex Set) algorithm [10], [21], governed by the updating rule

\[
x_{n+1} = P_{C_{\text{supp}}} \left( P_{C_{\text{dual}}} (x_n) \right).
\]

\(^1\)To determine whether \( C \) is non-empty is a nontrivial task and investigating this set is planned for future work. In the experiments conducted so far, the support constraints and redundancy were determined heuristically.

A. Compactly supported duals by truncation

In [19], Strohmer proposed a simple algorithm for the computation of compactly supported dual windows, which we will call the truncation method. Strohmer proposed to truncate the Wexler-Raz equations as described in the previous section and then solve the resulting equation system by computing the Moore-Penrose inverse, obtaining the least-squares solution. While the resulting windows satisfy the duality conditions, they are not very smooth and indeed show some discontinuity-like behavior, see also Figure 1(e,f). One of the goals of this contribution is the improvement of these undesirable effects.

V. NUMERICAL RESULTS

We present the construction of a smooth dual Gabor window with short support. Our setup considers \( g(\tau, 30, 60) \), i.e. a system with redundancy 2, where \( g \) is a “Nuttall” window [14] of length \( L_g = 120 \) samples, see Figure 1(a,b).

We aim at computing a dual that is supported on the same interval as the analysis prototype, yielding \( C_{\text{supp}} = \{ x \in \mathbb{R}^L : \text{supp}(x) \subseteq \text{supp}(g) \} \). To further provide reasonable localization and smoothness, we select the regularization functions \( f_1 = \| \cdot \|_1, f_2 = \| \mathcal{F}(\cdot) \|_1, f_3 = \| \nabla(\cdot) \|_2^2 \) and \( f_4 = \| \nabla \mathcal{F}(\cdot) \|_2^2 \). The result shown in Figure 1(c,d) shows the optimal dual window with regards to the regularization parameters \( \lambda_1 = \lambda_2 = 0.001 \) and \( \lambda_3 = \lambda_4 = 1 \). Those values are chosen experimentally by considering that they are balancing the effect of the regularization functions as described in Table I. As a reference, we included the least-squares solution provided by the truncation method, see Figure 1(e,f).

Minimizing the selected regularization functions improves upon the desired features, in particular smoothness (or frequency localization) with \( f_1 \) and time localization with \( f_2 \). The functions \( f_1 \) and \( f_2 \) avoid the solution to have a “M-shape”, i.e. multiple peaks. This is unwanted as the temporal or frequency positions becomes ambiguous. Indeed, minimizing the \( l^1 \)-norm will push all big coefficients to similar values.

The solution provided is assumed to perform perfect reconstruction on any signal with admissible length greater or equal to \( L \). More precisely, by [11, Eq. (60)], the maximum relative reconstruction error can be shown to be of the order of the precision of the machine, more precisely at \( 4.5 \times 10^{-14} \).

Simulations were performed using the LTFT [18] and the UNLocBoX matlab toolbox. A reproducible research addendum is available in http://unlocbox.sourceforge.net/tr/gwdwcuo.

In the experiment above, we constructed a smooth, well localized dual window, compactly supported with \( L_h = 120 \).

Considering the painless case, to guarantee the canonical dual window to be supported on \( \hat{g} = g \), enforces \( M \geq 120 \) therefore increasing the redundancy twofold, an unwanted side effect. Alternatively, in this setting, we could decide to keep the parameters \( a = 30, M = 60 \) fixed, but decrease the window size to \( L_g \leq 60 \). However, this construction provides a system with a more than 8 times larger frame bound ratio. Consequently, the resulting canonical dual window \( \hat{g} \), shown in Figure 2, shows bad frequency behavior and an undesirable, M-like shape in time. In contrast, the method proposed in
this manuscript allows the use of nicely shaped, compactly supported dual Gabor windows at low redundancies, without the strong restrictions of the painless case.

Our method can be applied in various situations to construct dual frames with properties more important for application than minimal $\ell^2$-norm. Future work will further be concerned with applying the findings herein to frames with a different structure, e.g. nonstationary Gabor frames [3].

Acknowledgement

This work was supported by the Austrian Science Fund (FWF) START-project FLAME (“Frames and Linear Operators for Acoustical Modeling and Parameter Estimation”; Y 551-N13).

We would like to thank David Shuman for his useful comments about the paper.

REFERENCES


