Semi-supervised Convolutive NMF for Automatic Piano Transcription

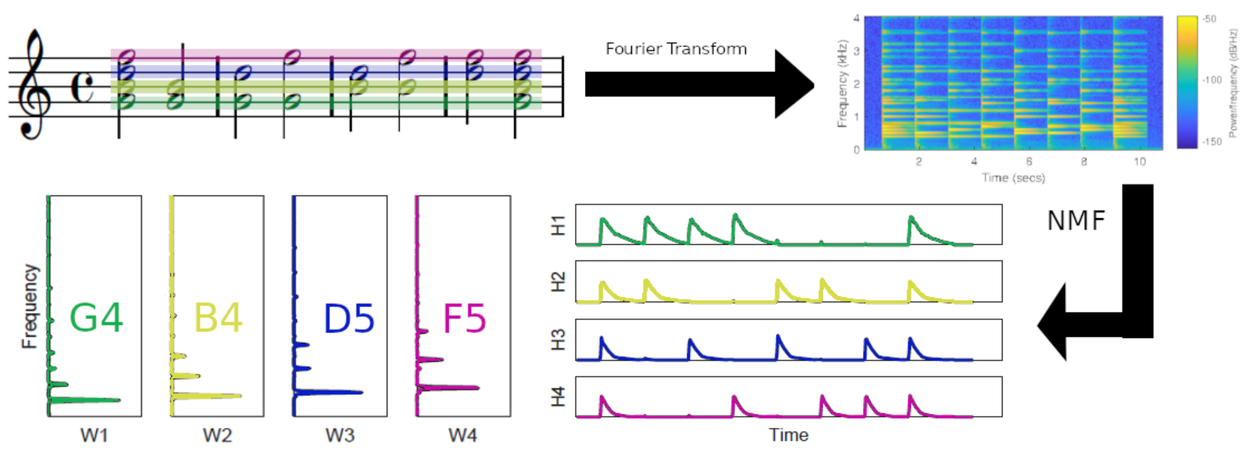
Haoran Wu, Axel Marmoret, **Jeremy E. Cohen**

IRISA, Rennes, France
CREATIS, CNRS, Villeurbanne, France

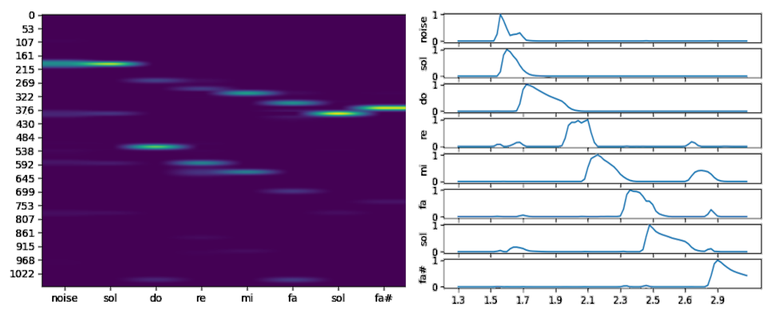
Paper link

Code link

Transcription principle with Nonnegative Matrix Factorization



NMF on simple audio

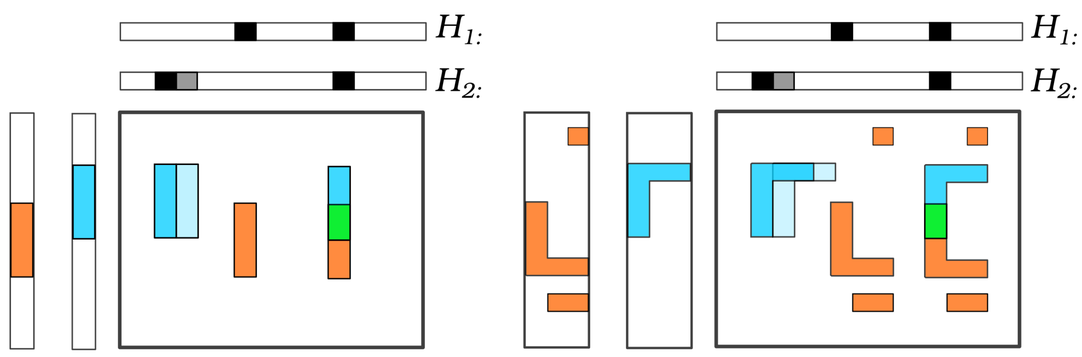


First 5 seconds of Jordu, transcribed

Challenges

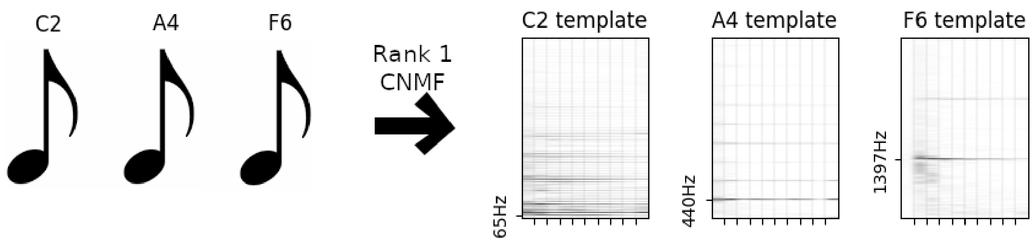
- More than one spectral template per note
- Time-dependent spectral templates
- Supervised but frugal

Proposed approach: Convolutive NMF



$$W, H \in \underset{W \geq 0,\, H \geq 0}{argmin}\ KL(Y, \sum_{q=1}^{r} W_{::q} * H_{q:})$$

Learning note templates...



Requires only individual notes recordings.

...then transcribing easily as a convex program

$$H \in \underset{H \geq 0}{argmin}\ KL(Y, \sum_{q=1}^{r} W_{::q} * H_{q:})$$

with W fixed.

Results on MAPS

			EN1		EN2		AkB1		AkB2		AkC		AkS		Sp		St	
thresh		τ	F	A	F	A	F	A	F	A	F	A	F	A	F	A	F	A
global	CNMF	5	78	65	70	55	88	80	75	62	83	72	80	69	81	70	75	61
		10	**85**	**75**	**77**	**64**	93	88	87	78	91	84	**88**	**79**	89	82	84	74
		20	83	72	76	63	**94**	**89**	**87**	**79**	**92**	**86**	87	79	**90**	**83**	**86**	**77**
	AD*		81	69	68	53	66	50	71	56	60	43	67	51	64	47	67	50
song	CNMF	5	82	70	74	59	90	82	80	69	87	78	84	74	86	77	81	69
		10	**88**	**79**	**80**	**68**	**95**	**91**	**90**	**83**	94	89	**90**	**82**	**93**	**87**	89	80
		20	85	75	78	66	**95**	**91**	**90**	**83**	**94**	**90**	89	81	92	87	**89**	**81**
	AD*		82	70	69	54	68	52	73	59	61	45	69	54	66	50	70	54
AD [3]			82	70	-	-	-	-	-	-	85	74	-	-	-	-	-	-
ByteDance DNN [8]			**89**	**81**	77	65	**98**	**97**	**95**	**90**	**98**	**96**	87	77	**97**	**95**	**95**	**90**

(Activations post-processing uses per-song threshold detector or global threshold detector)

Conclusions

- ✔ Comparable results with fully trained DNN
- ✔ Applicable when instrument in available, no registration
- ✔ Does not generalize well
- ✔ Time consuming

Some Relevent References

- **Seminal papers:** Smargadis2003 (NMF), Smaragdis2006 (CNMF)
- **Fully supervised DNN:** Cogliati2015,Sigtia2016,Hawthorne2018+2019,Kong2020,Shibata2021,Yan2021
- **Several templates per note:** Cheng2015, Ewert2016, Ewert2017
- **Time-dependent frequency templates:** Hennequin2010, Cheng2016, Gao2017

jeremy.cohen@cnrs.fr