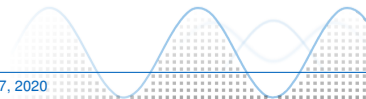


# Scene Detection in Skiing Videos

Niklas Walther

April 17, 2020



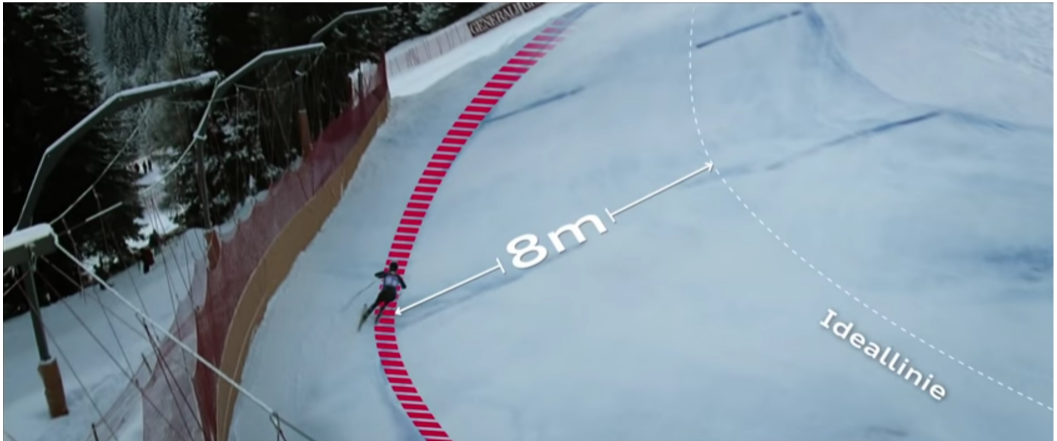
## Background: Alpine Ski Racing



# What makes an Alpine Ski Champion? Kinetics?



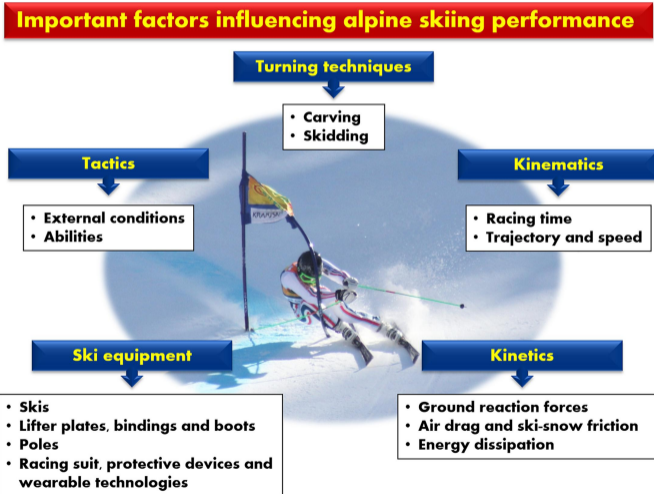
## The best Trajectory?



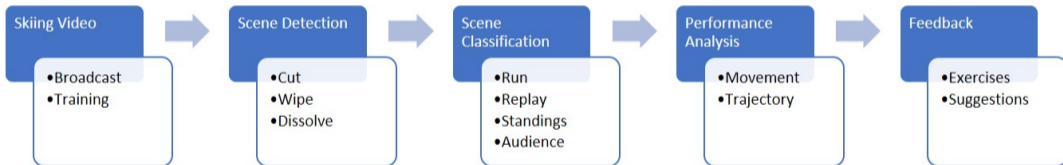
## Timing of Joint Movement?



# Guiding Model for Performance Analysis

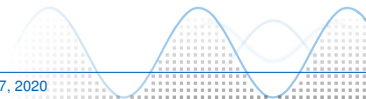


# Artificial Assistant



# Objective

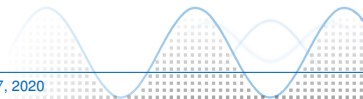
Developing an Application capable of detecting Scene Cuts in broadcasted Ski Races.



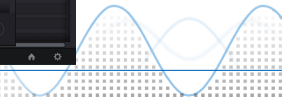
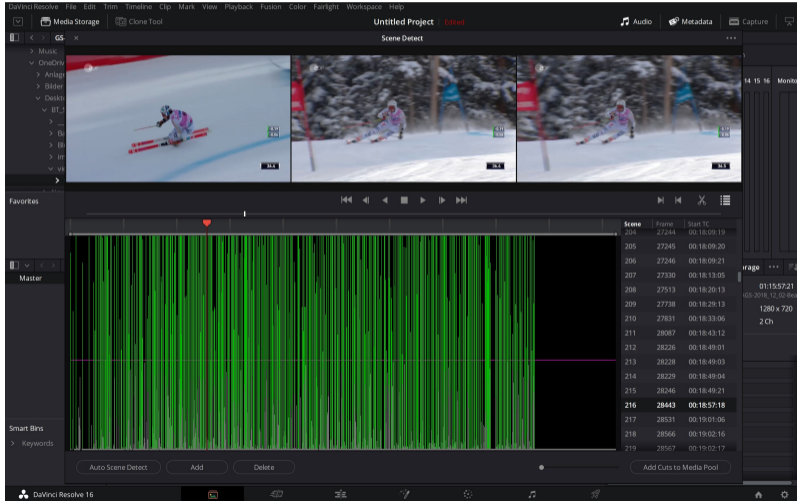


# Table of Contents

- 1 Background, Motivation & Objective
- 2 State of the Art
- 3 Cut Detection
- 4 Metric
- 5 Results
- 6 Conclusion



# State of the Art



## Direct Cut: Instant Frame Replacement between two Shots



(a) Frame 28442

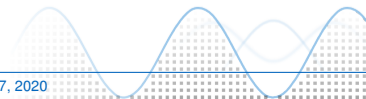


(b) Frame 28443



(c) Frame 28444

**Figure:** Cut between (b) and (c): The shots are taken simultaneously as also shown by the time overlay.



## FIS Wipe: Covering old with new Shot via FIS Emblem



(a) Frame 29768



(b) Frame 29772



(c) Frame 29780



(d) Frame 29782



(e) Frame 29786



(f) Frame 29792

Figure: The FIS Logo slowly slices away the old Shot.

# Cross Dissolve: Gradual Transformation by changing each Shot's Visibility



(a) Frame 9934



(b) Frame 9936



(c) Frame 9938



(d) Frame 9942

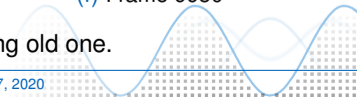


(e) Frame 9946

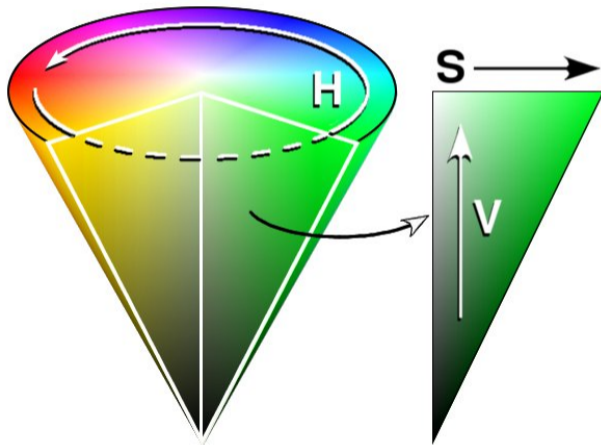


(f) Frame 9950

**Figure:** The new Shot seems to appear behind the fading old one.



## HSV Colour Space



Per Frame:

1. Take H, S & V Pixel Average.
2. Get Content Value as Mean of average Values.
3. Calculate Content Value Difference to prior Frame.

# Cut Detection

1. Definition of Ground Truth
2. Calculation of HSV Content Values per Frame
3. Threshold Value Estimation
4. Quality Measure for Threshold Values



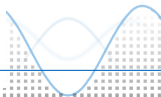
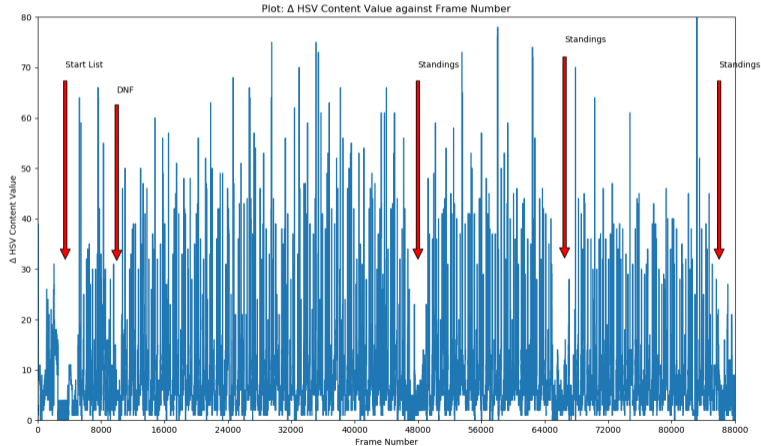
# Ground Truth Table

- 495 Transitions
  - ▶ 307 Cuts
  - ▶ 62 Dissolves
  - ▶ 32 Wipes
- Labels:
  - ▶ Frame Number
  - ▶ Time Code
  - ▶ Bib Number
  - ▶ Section
  - ▶ Notes

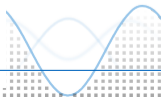
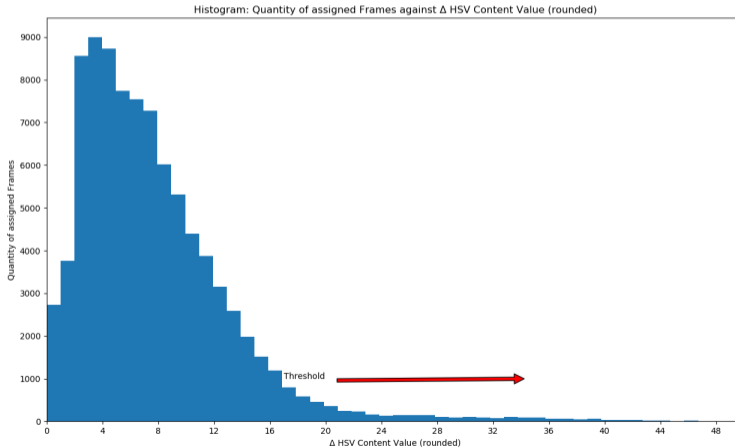
Scene Number	Start Frame	End Frame	diff	Start Time	End Time	diff	Start Number	Section	Note
1	1	64	64	0,04	2,56	2,52		Intro	
2	65	529	465	2,60	21,16	18,56		Gate 50 to 55	
3	530	545	16	21,20	21,80	0,60		Dissolve	
4	546	1.164	619	21,84	46,56	24,72		Forest	
5	1.165	1.169	5	46,60	46,76	0,16		Dissolve	
6	1.170	2.016	847	46,80	80,64	33,84		Start Preparation	
7	2.017	2.032	16	80,68	81,28	0,60		Dissolve	
8	2.033	2.559	527	81,32	102,36	21,04		Stands	
9	2.560	2.571	12	102,40	102,84	0,44		Dissolve	
10	2.572	3.866	1.295	102,88	154,64	51,76			Start List
11	3.867	3.885	19	154,68	155,40	0,72		Dissolve	
12	3.886	4.367	482	155,44	174,68	19,24		Gate 50 to 55	
13	4.368	4.378	11	174,72	175,12	0,40		Dissolve	
14	4.379	4.807	429	175,16	192,28	17,12		Gate 07 to 19	
15	4.808	4.820	13	192,32	192,80	0,48		Dissolve	
16	4.821	5.092	272	192,84	203,68	10,84		Start House	
17	5.093	5.104	12	203,72	204,16	0,44		Dissolve	
18	5.105	5.281	177	204,20	211,24	7,04		Start Preparation	
19	5.282	5.493	212	211,28	219,72	8,44		Start House	
20	5.494	5.841	348	219,76	233,64	13,88	1	Gate 01 to 07	
21	5.842	6.174	333	233,68	246,96	13,28	1	Gate 07 to 19	
22	6.175	6.343	169	247,00	253,72	6,72	1	Gate 19 to 25	
23	6.344	6.533	190	253,76	261,32	7,56	1	Gate 25 to 31	
24	6.534	6.660	127	261,36	266,40	5,04	1	Gate 31 to 35	
25	6.661	6.948	288	266,44	277,92	11,48	1	Gate 35 to 43	
26	6.949	7.177	229	277,96	287,08	9,12	1	Gate 43 to 50	
27	7.178	7.328	151	287,12	293,12	6,00	1	Gate 50 to 55	
28	7.329	7.546	218	293,16	301,84	8,68	1	Gate 55 to 60	
29	7.547	7.640	94	301,88	305,60	3,72		Finish	
30	7.641	7.690	50	305,64	307,60	1,96		Stands	
31	7.691	7.790	100	307,64	311,60	3,96		Finish	
32	7.791	7.814	24	311,64	312,56	0,92		FIS Wipe	
33	7.815	7.997	183	312,60	319,88	7,28			Repetition
34	7.998	8.018	21	319,92	320,72	0,80		FIS Wipe	
35	8.019	8.079	61	320,76	323,16	2,40		Finish	



# HSV Content Value per Frame



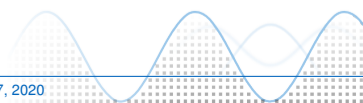
# HSV Content Value Distribution



# Metric

		Ground Truth Transition	
		Positive	Negative
Transition in Detection	Positive	C	F
	Negative	M	N

- Correct Detection (C)
- False Detection (F)
- Missed Detection (M)
- No Detection (N)



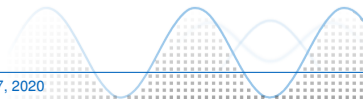
## Metric: Accuracy (A)

		Ground Truth Transition	
		Positive	Negative
Transition in Detection	Positive	C	F
	Negative	M	N

$$A = \frac{C + N}{C + M + F + N} \quad (1)$$

$$A = \frac{0 + 0.995}{0 + 0.005 + 0 + 0.995} \quad (2)$$

$$A \rightarrow 99.5\% \quad (3)$$



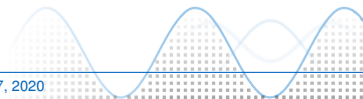
## Metric: Recall (R)

		Ground Truth Transition	
		Positive	Negative
Transition in Detection	Positive	C	F
	Negative	M	N

$$R = \frac{C}{C + M} \quad (4)$$

$$R = \frac{1}{1 + 0} \quad (5)$$

$$R \rightarrow 100\% \quad (6)$$



## Metric: Precision (P)

		Ground Truth Transition	
		Positive	Negative
Transition in Detection	Positive	C	F
	Negative	M	N

$$P = \frac{C}{C + F} \quad (7)$$

$$P = \frac{1}{1 + 0} \quad (8)$$

$$P \rightarrow 100\% \quad (9)$$

## Metric: Quality (Q)

		Ground Truth Transition	
		Positive	Negative
Transition in Detection	Positive	C	F
	Negative	M	N

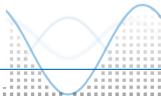
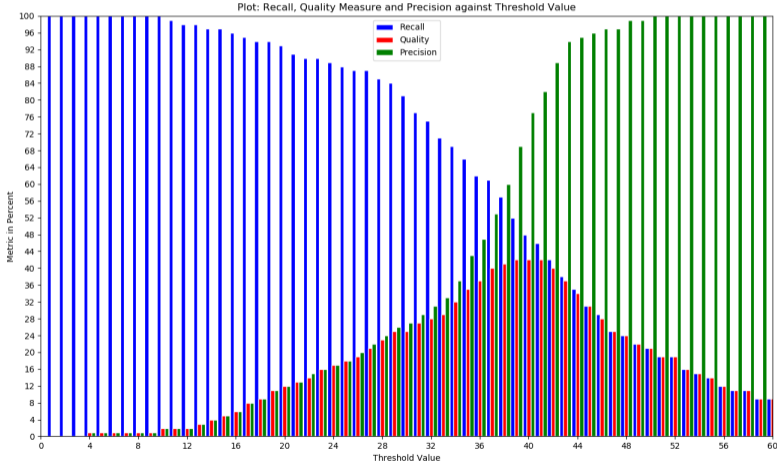
$$Q = \frac{C}{C + M + F} \quad (10)$$

$$F1 = \frac{2 \cdot P \cdot R}{P + R} \quad (11)$$

$$\left( = \frac{2 \cdot C}{2 \cdot C + M + F} \right) \quad (12)$$



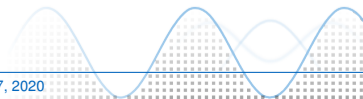
# Quality Measure by Threshold



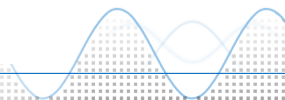
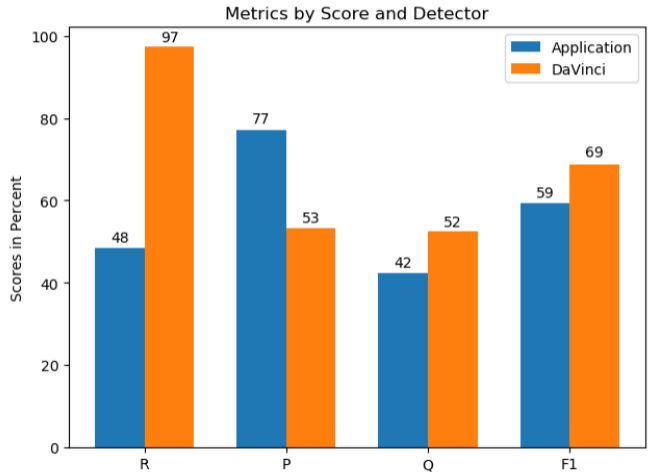


## Results

	Application	Da Vinci Resolve
Processing Time	ca. 90min	ca. 10min
Detects Cuts?	✓	✓
Is automatable?	✓	✗
Is extendable?	✓	✗

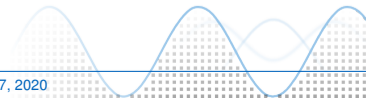
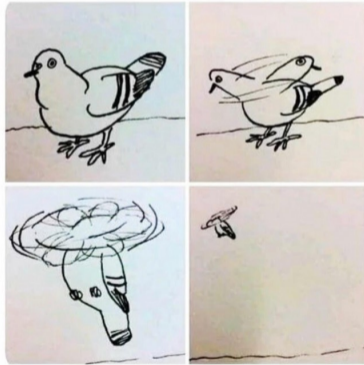


# Results



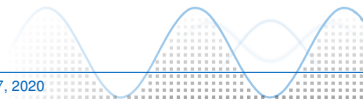
# Conclusion

When your program  
is a complete mess,  
but it does its job

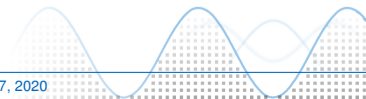
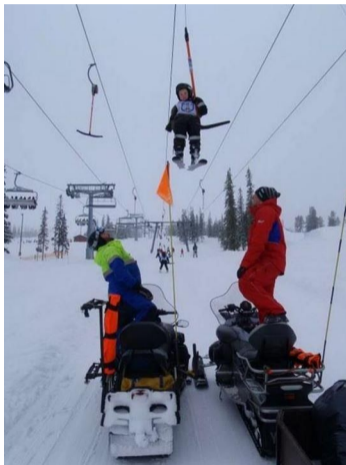


## Conclusion

- The developed Application is capable of detecting Scene Cuts in broadcasted Ski Races. → Objective fulfilled.
  - The Application is outperformed by Da Vinci Resolve in terms of Processing Time.
  - While competitive Quality wise, the Auto Scene Detection slightly surpasses the Application regarding the pure Cut Detection.
  - When outsourcing Scene Detection, the Artificial Assistant is not fully automatable.
- ⇒ With the scripted Parsing Tools and additional manual input during Processing, Blackmagic Design's Solution can be used for the Artificial Assistant's Scene Detection.



Thank you for your Attention.



## Image References

- <https://www.youtube.com/watch?v=EDKHfgsd320>, last accessed on 16/04/20
- M. Supej H.-C. Holmberg, Determinants of Skiing Performance in *Frontiers in Physiology* (2019), <https://www.frontiersin.org/articles/10.3389/fphys.2019.00111/full>, last accessed on 15/04/20
- Eric Pierce, HSV Colour Space in Wikipedia (2005), [https://de.wikipedia.org/wiki/HSV-Farbraum#/media/Datei:HSV\\_cone.jpg](https://de.wikipedia.org/wiki/HSV-Farbraum#/media/Datei:HSV_cone.jpg), last accessed on 15/04/20
- <https://devhumor.com/media/when-your-program-is-a-complete-mess>, last accessed on 16/04/20
- <https://i.pinimg.com/originals/54/77/c5/5477c53d9457b4211d9bcb8761750158.jpg>, last accessed on 16/04/20