

You Tube Video Genres. Amateur *How-To* Videos versus Professional Tutorials

Andreea Mogoş¹, Constantin Trofin²

Abstract: In spite of the fact that there is a vast literature on traditional textual and visual genre classifications, the categorization of web content is still a difficult task, because this medium is fluid, unstable and fast-paced on one hand and, on the other hand, the genre classifications are socially constructed through the tagging process and the interactions (commenting, rating, chatting). This paper focuses on YouTube tutorials and aims to compare video tutorials produced by professionals with amateur video tutorials.

Keywords: genre classification; how-to-do-it features; video tutorials; YouTube

1. Genre Classification Issues

Florea & Catarig (2011, p. 96) defined print media genres as “textual-discursive forms of information mise-en-scène”. For descriptive purposes, the researchers combined two sets of criteria, proposed by Lochard (1996) and Charaudeau (1997). Thus, two levels were identified as genre-defining: *the pragmatics of discourse* (discursive modes; macrostructural organization) and *the textual and enunciative configuration* (structure and length of the text; role of the enunciative instance in the text; microstructure – syntax and stylistics). When discussing about traditional media genres, one can use several existing typologies in order to identify a genre. Due to their formal specific features, genre categories have both epistemological and functional dimensions: they are ways of organizing and defining content, but also ways of organizing social actions (Bawarshi, 2001).

¹ Associate Professor, PhD, Babeş-Bolyai University, Romania, Address: 1 Kogălniceanu Str., Cluj-Napoca, Romania, Tel: +40-264-40.53.00, Fax: +40-264-59.19.06, E-mail: mogos@fspac.ro.

² Assistant Professor, PhD, Babeş-Bolyai University, Romania, Address: 1 Kogălniceanu Str., Cluj-Napoca, Romania, Tel: +40-264-40.53.00, Fax: +40-264-59.19.06, Corresponding author: trofin@fspac.ro.

Traditional media still relies on standard media genres (such as news story, feature, interview, column, investigative report), but the web creates an environment where message producers are not only the professionals invested with authority, who are aware of rules and norms that are governing the media communication process, but the members of the audience, too.

“Functionally defined supergenres are medially stable and have relatively clear genre antecedents in pre-digital discourse, whereas formally defined subcategories are more diverse, more prone to change, and may easily be generated within a changing medial environment.” (Heyd, 2009, p. 241)

Therefore, when dealing with genres of Web pages, there are two important aspects to be taken into account (Santini, 2007): on the one hand, the Web is fluid, unstable and fast-paced; on the other hand, genres on the Web are instantiated in Web pages, which are a complex type of document, more composite and unpredictable than paper documents. These two aspects are interwoven and often result in classification difficulties that could be analyzed in terms of two broad textual phenomena: *genre hybridism* and *individualization*. Digital genre candidates are hybrid or “bridging” genres (Herring et al., 2005) because they are medially migrating genre instantiations that are typically based on antecedents in the traditional spoken or written media.

When talking about online videos, one should take into account that the diversity of these genres arises from the multiplicity of their contexts of use. Therefore, categorization of online videos is often treated as a *tag* suggestion task; tags can be generated by individuals or by machine classification. Yew & Churchill (2011) suggest that categorization can be determined socially, based on people’s interactions around media content without recourse to metadata that are intrinsic to the media object itself:

“The social consumption of media can alter the way content is perceived and categorized. Media content that we believe to fit a particular genre is both constituted by, and constitutive of, the changing social contexts in which that content is produced, shared and consumed; genres are socially constructed.”

Every seventh human being is a YouTube user (this is the equivalent of more than 1 billion people). 300 hours of video are uploaded to YouTube every minute and everyday people watch hundreds of millions of hours on YouTube and generate billions of views.¹ In addition to simply watching and sharing videos on YouTube, people interact around the video by leaving comments, rating content or, in the case

¹ <https://www.youtube.com/yt/press/statistics.html>.

of synchronous sharing, by chatting. This leads to the idea that often YouTube video genres categorization is socially determined by a complex mechanism and multiple criteria: the tags added by the user; the verbal/visual interaction through video answers, comments, ratings, and chat. Given the huge amount of videos available and constantly uploaded, automatic categorization methods were necessary in order to be able to deal with this type of content.

Over the past years an increasing literature trying to bridge the gap between the human perception of genre and the automatic categorization of genre in classifying online videos emerged. After reviewing the literature on methods, Brezeale & Cook (2008) identified four groups of methods for performing automatic classification of video: *text-based approaches*, *audio-based approaches*, *visual-based approaches*, and those that used some *combination of text, audio, and visual features*. Most authors incorporated a variety of features into their approach, in some cases from more than one modality. Thus Wang, Xu, & Chng (2006) proposed a multi-level framework to automatically recognize the genre of the sports video, consisting of a classifier using low-level visual/audio features to evaluate the video clips; Wang et al. (2010) proposed an approach that combines multiple data sources for wild YouTube video categorization¹; Ionescu et al. (2012) proposed three categories of content descriptors, extracted at temporal, color and structural level, in order to classify seven common video genres².

2. From how-to Feature Stories to Video Tutorials

The paper will briefly review the main categories of the *feature* genre and their characteristics with the purpose to establish if *how-to-do-it* articles and video tutorials share common traits. The major types of feature articles used by magazines, newspapers, newsletters, and online publications (Garrison, 2004, p. 8) are: “*color stories that emphasize descriptive writing, human interest stories, personality*

¹ The authors proposed a fusion framework in which each data source is first combined with the manually-labeled set independently. Then, using the hierarchical taxonomy of the categories, a Conditional Random Field (CRF) based fusion strategy is designed. Based on the final fused classifier, category labels are predicted for the new videos.

² According to the authors, at temporal level, video content is described with visual rhythm, action content and amount of gradual transitions. Colors are globally described with statistics of color distribution, elementary hues, color properties and relationship. Finally, structural information is extracted at image level and histograms are built to describe contour segments and their relations. The proposed parameters are used to classify seven common video genres, namely: animated movies/cartoons, commercials, documentaries, movies, music clips, news and sports.

sketches and profiles, seasonal feature stories, reviews and critical feature writing, aftermath and follow-up feature stories, the feature series, travel article writing, personal experience stories, how-to-do-it articles, humor writing, and technical and specialized features.”

Feature stories focus on home how-to-do-it subjects, consumer and shopping ideas, profiles, health and medicine suggestions, child-care news, gardening, cooking and food ideas, and much more. In some newspapers, entertainment news is part of lifestyle coverage, but this category of news creates a different set of feature needs for newspapers. Writers specializing in entertainment provide features on individuals, reviews of their performances, and insight into their creative activities.

More specifically, the how-to article explains how something is made, built, cooked, protected, purchased, or otherwise accomplished by an expert on the subject. These articles are often found in home and garden sections, food and cooking sections, and increasingly so in consumer-based sections of newspapers and in the similarly named departments of magazines or newsletters. How-to articles are the most frequently published forms of service article. Alexander (1975, p. 213) explains why the how-to article is a popular story form with editors: *“In a pragmatic nation, Americans look more and more to magazines to advise them and show them how to do the things that are important in their lives”*.

Alexandre’s 40 years old argument could be reformulated, taking into account that newspapers and magazines are decreasing in popularity, while the online content becomes the main source of information, entertainment and propaganda. The same phenomenon also regards television, which is constantly losing ground to online video. The world itself is changing, and the public changes along with it. More and more viewers are leaving old, traditional media behind, embracing the full availability and easy access on practically anything on the internet, anytime. And if you cannot find the content you are looking for, you can now produce it yourself, for the benefit of others searching for the same type of content. Therefore, there is a striking difference between how traditional television and online videos (most specifically the YouTube content) approach the viewer:

- TV ratings are generated by the joint quality of **content** and **production** which add up to an attractive visual product, achieving a market rank of a certain TV station, which, if wisely programmed, as a consequence will produce a better GRP (Gross Rating Point)¹ which finally would mean more

¹ Gross Rating Points (GRPs) equal Reach times Frequency, expressed as a percentage. <http://www.marketing-metrics-made-simple.com/gross-rating-points.html> (June 26, 2015).

money gained on a competing market environment. Viewers would watch a program because of its content and, to a large extent, because of its appearance (high quality video and audio). Most viewers watch the whole of the show.

- the number of views of tutorials on Youtube.com is rather related to **content** than appearance. People watch a tutorial to find a quick solution to a specific demand or problem. So the clip which provides the swiftest, most direct and easiest to understand explanation is likely to get the most views, regardless of how well it's shot, edited or the quality of sound, as long as the message gets through. Most people don't even watch the entire clip. They go directly to their point of interest.

There are incredibly numerous tutorial films or clips on YouTube.com. Searching "how to" produces about 155 million results, while searching "tutorial" produces almost 83 million results.

There are tutorials about pretty much everything one can think of. The most popular tutorials are about beauty care: Make-up, finger nails, a.s.o. Yet, the most popular of all (17, 910, 415 views on September 23, 2015) is the "Severed fingers tutorial"¹, which teaches viewers makeup special effects to create the illusion they cut off their fingers with a large knife, which shows more curiosity than real interest. Just like in television, but this is a matter of human nature.

The most viewed "how to" film got 12,222,167 views by September 23, 2015. The film is called "10 HOW TO PRANKS for Home and Office"² and teaches viewers, naturally, how to set up pranks. Again, explained by the human nature and a certain perception on what is funny.

All in all, tutorials teach people to solve different situations or find answers to everyday questions. Tutorials are not all done the same way though, or in the same quality range. Some are professionally produced while others look rather amateurish. Yet this is not a viewing criterion. This is a clear indicator of the fact that internet rules do not apply in television and vice-versa, due to the different nature of the viewing purpose and the viewers themselves.

¹ <https://www.youtube.com/watch?v=2P9rE9hoOZg> , published on Aug 5, 2014 by Freakmo.

² <https://www.youtube.com/watch?v=Ok-susL-OpY> , published on Jun 13, 2014 by howtoPRANKitup.

3. Professional Tutorials vs. User Generated Content

Professional videos are generally produced (or ordered) by manufacturing companies or servicing companies in order to help owners to operate, repair or improve the efficiency of their products and generally follow the same production patterns as advertisements. User generated clips basically aim the same goals, but they are produced by passionate users or craftsmen who promote their businesses.

3.1. Methodology

The analysis focuses on describing two aspects of YouTube video tutorials (production instances and content characteristics) which enable a comparison between professional and amateur YouTube tutorials.

The paper will answer a set of questions that could offer a better understanding of the video tutorials **production**: *when was it made? where was it made? who made it? was it made for someone else? what technologies does its production depend on?*

Further on, the research will focus on **content** characteristics by answering the following questions: *what is being shown? what are the components of the image? How are they arranged? is it one of series? where is the viewer's eye drawn to in the image, and why? what is the vantage point of the image? what relationships are established between the components of the image visually? how has its technology affected the text?*

3.2. Analysis

There is an important thing that must be thoroughly understood: The most important feature of a YouTube tutorial consists in its *utility*. They are always being issued on consumer demand. While professionally produced tutorials meet a demand foreseen by the manufacturer of a product or the supplier of a certain service in order *to enhance sales or aftermarket support*, user generated tutorials appear when a certain user encounters a problem regarding a product in his current usage. Whether they find solutions on their own or with the help of a professional who in most cases is an acquaintance, this is a way of sharing useful personal experience with others who might encounter the same problem, in the spirit of YouTube, a very popular tool and a democratic environment.

So, even though they are conveyed by the same carrier, YouTube, and their final goal is similar, there are a number of striking **differences** between the “professional” and the “profane” approach, which are going to be summarized as follows:

The professional approach involves all the rules and rigors of video (and television) production. Carefully planned high definition shooting in terms of both composition and exposure, proper lighting for each shot, lots of details, preplanned and motivated camera movements, a balanced and dynamic edit, fair proportions and professionally mastered video graphics. Action is being carried out by carefully selected actors, following a good and well developed line of story based on a professional script. Sound is very carefully taken care of, with well-engineered recording of ambient sound, professional commentary voices recorded in studio conditions, original musical score and serious post production treatment. All the above mentioned are professionally dealt with even if the clip teaches you how to change the oil of your lawn mower.

The user generated tutorials widely differ, based on the video training and skills of the producer, who sometimes does not even have a YouTube channel. Clips range from somehow professionally looking clips to rather home video ventures. The latter have a very raw approach, with no regard for the knowledge of visual production. They are usually one shot or cold cut films, with the camera placed on a tripod or even handheld by a friend who occasionally follow the action, when it's not locked on a wide shot. Ambient light is usually the only light used for shooting. Sound is usually picked up by the built-in microphone of the video camera and, in some cases, the producers feel the urge to put in some music. The “actor” is usually the “producer”. He speaks freely, improvising on the way. Another way of doing it is continuous point of view shooting, with “live” commentary from behind the camera, since the actor, producer and camera operator are the same person. It is similar to home video. Is that a problem? Not really, if the message gets through and the clip provides a useful solution to a specific problem. The clip mentioned earlier, that gathered almost 18 million views, was produced in the “profane” way. Such a tutorial can be produced and uploaded in less than an hour, while a professional tutorial takes time to produce, as it undergoes all stages of professional pre-production, production and post-production. Professional tutorials are timed to be released simultaneously with the launch of a new product or service. While professional tutorials are treated as film productions, most user generated tutorials are merely home videos. These different approaches affect accordingly the technology and costs implied. Professional cameras, lighting, sound capture and edit, carefully selected sets and

actors, as opposed to consumer handy cams with built-in microphones shooting in environmental light with no set preparation whatsoever, with the producer also impersonating the “actor”. In most cases, professional tutorials are being staged and shot in studios, where the sets are built to resemble a workshop, a construction site, a home or a garden. User generated tutorials are shot in actual workshops, construction sites, homes, gardens and in a significant number of cases, in personal garages. Many professional tutorials have a cinematographic approach, sometimes presenting a story and well defined characters. In most cases they are multi-camera productions, sometimes, if action movement is implied, with the use of cranes, jibs, dollies, steady cams or speed lines. Action shooting and action edit add to a cinematic tension, meant to add up to a sense that the viewer is watching a quality movie, transferring its grandeur and seriousness to the product it presents, very much as in the case of expensive TV commercials. Consumer videos don’t present this approach for a simple reason: they are not intended to praise a certain product. In most cases, people bought a certain product by chance, they had encountered a problem and faced it successfully and they want to share their experience with others. So, in two easy steps, they first present the problem and then they solve it. In many cases, the whole process is shown in a single shot as in any home video, sometimes inserting some captions, but no professional edit or even framing are usually taken into consideration. In some cases, products even get heavily criticized. But as long as someone paid for that product, it must be put to work.

So the “producers” could not care less about technicalities of film making, such as sets, lights, edit or acting. There is no direct connection in their mind between replacing the blade of a mechanical saw and the light on the background. There are, though, situations in which certain technicians promote their own businesses, even offering to produce other tutorials on demand. In such cases a little more attention is shown for the “looks” of the video, as well as the presenter. In terms of costs, it is clear that a video production sponsored by a manufacturing company for instance, will look professional, as in most cases they are a part of expensive company promotion projects. Consumer videos practically cost nothing, as they are not meant to support anything but fellow citizens in need, and in some cases, support small businesses such as repair shops or beauty salons.

Content also differ. In many cases, professional tutorials, based on a script, create a story that the viewer could easily identify with, to introduce the situation to be solved, to “draw the viewer in” before presenting the situation. In many cases the producers create series of tutorials with the same production patterns, same actors

and, of course, same visual approaches. Of course, the users often produce series of tutorials but usually the only element of continuity is the presenter and sometimes, the garage. Probably the answer to the question: “Does it matter?”, the producers’ answer would be : “Who cares?”

Table 1. Professional how-to videos vs. amateur tutorials. A comparison

	Professional <i>how-to</i> videos	Amateur tutorials
Preproduction	Script, set design, lighting and sound design, talent casting, directing, rehearsals, archive materials, specialized supervising.	Unscripted and undirected. Set and camera setup in available facilities.
Production	Carefully planned shooting, multi-camera shootings. Specialized camera support systems. Preplanned composition, exposure, lighting, sound and talent movement.	Random shooting, small number of shots, frequently single shot videos, adapted to set and lighting conditions. In some cases, hand held shooting, without any camera support.
Postproduction	Elaborated edit and graphics obeying rules of filmmaking.	Raw or no edit, sometimes unsophisticated graphics or additional music soundtracks.
Costs	High costs based on a structured budget.	Low or inexistent using available materials on site.
Content	Structured information delivery	Unscripted, direct speech.
Aim	To inform consumers about new products or services or to help them self-maintain products in order to turn viewers into steady customers of the product.	To share personal experience with other people interested in the matter.

4. Conclusions

Internet rules do not apply in television and vice-versa, due to the different nature of the viewing purpose and the viewers themselves.

As we earlier concluded, the essence of any YouTube tutorial consists in its utility. Their existence is determined by consumer demand. So under the given circumstances and considering the analysis performed in this paper, why would anyone bother investing in “professional” tutorials? Why not give up all you have learnt about composition, exposure, lighting, storytelling, scripting and other professional matters? Why not do it the “citizen” way? Because of those important things that drive all forms of journalism as well: credibility and trust. These can really make the difference, whether you are dealing with respected brands and businesses or you just want to be taken seriously. All in all, there’s not a real problem in producing tutorial videos the “citizen way”, as long as the “citizen” is an “educated citizen”, perceived as such by the viewers.

5. References

- Brezeale, D. & Cook, D. J. (2008). Automatic video classification: A survey of the literature. *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on*, vol. 38 (3), pp. 416-430.
- Florea, L.S.; Danciu, I.M. & Catarig, A. T. (2011). *Gen, text și discurs jurnalistic: tipologia și dinamica genurilor în presa scrisă română și franceză/Gender, text and journalistic discourse: the typology and gender dynamics in Romanian and French written press*. Bucharest: Tritonic.
- Garrison, B. (2009). *Professional feature writing*. London: Routledge.
- Herring, S. C.; Scheidt, L. A.; Wright, E. & Bonus, S. (2005). Weblogs as a bridging genre. *Information Technology & People*, vol. 18 (2), pp. 142-171.
- Heyd, T. (2009). A model for describing “new” and “old” properties of CMC genres. *Genres in the internet: Issues in the theory of genre*, pp. 239-262.
- Ionescu, Bogdan; Seyerlehner, Klaus; Rasche, Christoph; Vertan, Constantin & Lambert, Patrick. Content-Based Video Description for Automatic Video Genre Categorization, pp.51-62 in Schoeffmann, K., Merialdo, B., Hauptmann, A. G., Ngo, C. W., Andreopoulos, Y., & Breiteneder, C. (Eds.). (2012). *Advances in Multimedia Modeling: 18th International Conference, MMM 2012, Klagenfurt, Austria, January 4-6, 2012, Proceedings* (Vol. 7131). Springer.
- Rose, G. (2012). *Visual methodologies: An introduction to researching with visual materials*. London: Sage.
- Santini, M. (2007, January). Characterizing genres of web pages: Genre hybridism and individualization. In *System Sciences, 2007. HICSS 2007. 40th Annual Hawaii International Conference on* (pp. 71-71). IEEE.

Wang, J.; Xu, C. & Chng, E. (2006, August). Automatic sports video genre classification using pseudo-2d-hmm. In *Pattern Recognition, 2006. ICPR 2006. 18th International Conference on*, Vol. 4, pp. 778-781. IEEE.

Wang, Z.; Zhao, M.; Song, Y.; Kumar, S. & Li, B. (2010, June). Youtubecat: Learning to categorize wild web videos. In *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, pp. 879-886. IEEE.

Yew, J. & Shamma, D. A. (2011, February). Know your data: Understanding implicit usage versus explicit action in video content classification. In *IS&T/SPIE Electronic Imaging*, pp. 78811A-78811A. International Society for Optics and Photonics.

Yew, J., Shamma, D. A. & Churchill, E. F. (2011, May). Knowing funny: genre perception and categorization in social video sharing. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 297-306. ACM.